UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

February 19, 2020

EPA-CASAC-20-002

The Honorable Andrew R. Wheeler
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: CASAC Review of the EPA's *Integrated Science Assessment for Ozone and Related Photochemical Oxidants (External Review Draft – September 2019)*

Dear Administrator Wheeler:

The Chartered Clean Air Scientific Advisory Committee (CASAC) met on December 3-6, 2019, and on February 11-12, 2020, to peer review the EPA's *Integrated Science Assessment for Ozone and Related Photochemical Oxidants (External Review Draft – September 2019)*, hereafter referred to as the Draft Ozone ISA. The Chartered CASAC approved the report on February 11, 2020. The CASAC's consensus responses to the agency's charge questions and individual review comments from members of the CASAC are enclosed. Questions from CASAC members to a pool of non-CASAC member consultants and their responses are also enclosed. Major comments and recommendations are highlighted below and detailed in the consensus responses to charge questions.

On overarching process issues, the CASAC strongly recommends that the EPA consider restoring a traditional interactive discussion process in which the CASAC can interact directly with external expert panels, while also keeping the option of obtaining written responses from external experts to specific questions. The CASAC offers additional process recommendations in its review of the EPA's Draft Ozone Policy Assessment (PA).

Overall, the CASAC finds that the Draft Ozone ISA, while providing useful reviews of many aspects of ozone exposures and human health effects in selected studies, does not provide a comprehensive, systematic assessment of the available science relevant to understanding the public health impacts of changes in ambient concentrations of ozone. The CASAC recommends that the following key points be addressed in the final Ozone ISA:

- Critically review, synthesize, and discuss available scientific evidence on how changes in public health effects depend on changes in ambient ozone exposures. This is a crucial scientific topic for informing the Ozone PA and should be thoroughly addressed in the Ozone ISA.

- Clarify criteria used to select, evaluate, weight, and summarize studies; provide details of how the criteria were applied to individual studies and what the results were; and explain how key conclusions were derived from the results.
- Clarify the meaning and derivation of stated key causal conclusions. Causal determination judgments stated in the Draft Ozone ISA are ambiguous, and sometimes appear subjective and arbitrary. The meanings of the causal determination terms used should be specified (e.g., does "causal" refer to necessary causation, sufficient causation, or something else?) and how causal conclusions are reached from the evidence presented should be made explicit and transparent. The CASAC recommends that the EPA seek help from external experts in relevant areas, e.g., via the National Academies, to strengthen and clarify its framework for causal inference.

Turning to the parts of the Draft Ozone ISA, the CASAC finds that the Executive Summary provides a concise summary of key findings in the Draft Ozone ISA, but that the information summarized is unclear in essential respects and does not accurately represent the totality of available high-quality scientific evidence on health effects of changes in exposures to ambient ozone. The CASAC recommends adding the following information to the Executive Summary:

- Summarize available scientific evidence on how changes in public health effects depend on changes in ozone levels.
- Present summary results from a systematic review and critical evaluation and synthesis of relevant studies informative about public health effects of changes in ozone levels, including negative ones that have been omitted from the Draft Ozone ISA. As part of this review, discuss possible confounding (e.g., by region, season, month, year) in detail, and how it was or was not addressed.
- Discuss causal biological mechanisms of inflammation-related health effects.
- Summarize the results of comprehensive quantitative uncertainty and sensitivity analyses showing how conclusions change for plausible variations in assumptions, interpretations of terms, selection and weighting of studies, and judgments on which the conclusions depend.

The Integrated Synthesis, as well as the Executive Summary, should be revised to thoroughly address the preceding points, and also to clarify the treatment of wildfire contributions to ozone exposure, and their implications for the National Ambient Air Quality Standards (NAAQS). It should clarify to what extent ozone-associated physiological effects represent adverse health effects. It should provide fuller and more accurate coverage of relevant scientific literature (including relevant methodological advances and international data) and more balanced coverage of negative studies and of literature on nonlinear effects.

Appendix 1 should be revised to provide a detailed discussion on the uncertainty associated with the emissions inventory (by pollutant and source sector); add national maps of county-level emissions of precursors, and discuss their relative importance for ozone formation; clarify the impact of inter-annual variability and longer term trends in meteorological effects on ozone design values; discuss topographical effects on meteorology, ozone formation, and ozone transport; describe ground-based ozone lidar instruments and satellite data; emphasize the importance of performing a comprehensive model performance evaluation when using regional chemical transport models (including for precursors); and clarify how exceptional events are accounted for in health studies and risk analyses. Appendix 1 should discuss the shifting of ozone peak concentrations from summer to spring and fall that is occurring in many parts of the country, along with trends in ozone precursors. It should compare

estimates from measurement-based and modeling-based approaches to understand differences and reduce uncertainty in U.S. Background (USB) ozone estimates and should discuss the ozone design values that can result from USB.

Appendix 2 should add information on the Air Pollution Exposure model (APEX) and Stochastic Human Exposure and Dose Simulation (SHEDS) models and further discuss ozone infiltration in vehicles, as well as providing more detailed discussions of the uncertainties and variability associated with the Consolidated Human Activity Database (CHAD), indoor-outdoor (I/O) ratios to describe infiltration of ambient ozone into homes and buildings, and personal exposure-ambient concentration (P/A) ratios. The final ISA should include more detailed discussions of summertime confounding of ozone effects by copollutants and the impact of exposure measurement error on effect estimates in epidemiology studies.

For Appendices 3-7, the CASAC recommends that the final ISA should more fully and explicitly address chance, bias, confounding, and other non-causal sources of associations in its analyses of epidemiology study quality; provide a more balanced and accurate summary of study results for each health endpoint (including available positive, negative, and null results); appropriately compare animal to human ozone doses when extrapolating animal exposures to potential human risks; present dose information in biological plausibility discussions; include exposure durations and exercise levels when presenting results, particularly for controlled human exposure (CHE) studies; clarify comparisons of responsiveness in people with and without preexisting conditions; clarify mean measured concentrations in summarizing study results; and more clearly distinguish between negative results (i.e., effects not detected in adequately powered studies) and absence of results (effects not looked for). The final ISA should discuss the scientific significance of conflicting and/or inconsistent evidence. For animal studies, it should further discuss what is currently known about no-effect and low-effect concentrations and comparability of animal models to human diseases. For epidemiological studies, the ISA should state that variability and error in the variables can linearize C-R functions and obscure thresholds whenever it concludes that epidemiological relationships between ozone and health effects appear to be linear no-threshold (LNT), and it should apply appropriate technical methods, including errors-in-variables methods, (and encourage the epidemiological community to apply them) to address this particular concern. If possible, the ISA should include these adjustments when applying epidemiology C-R functions to their risk assessments. The ISA should address the adversity and clinical significance of important health effects, such as changes in fasting blood glucose, and should ensure that all relevant information is included in the study figures and tables. The CASAC finds that Figure 3-1 provides a useful synthesis of known and suspected biological pathways mediating ozone respiratory health effects, but recommends several refinements and encourages the EPA to include both positive and negative studies, as well as information about exposure concentrations, in presenting biologically plausible pathways.

For the short-term ozone effects on metabolic endpoints, the data do not justify the “likely” causal determination. “Suggestive” appears to be a more appropriate designation. For the causality designation for long-term ozone effects on metabolic endpoints, the evidence does not justify the “likely” determination. Designation of ozone effects on fertility and reproduction as “suggestive of causality” is also not well supported by the available data. The CASAC recommends that these causality designations and their rationales be reconsidered in the final ISA. For short-term ozone exposure and mortality, and for short-term ozone exposure and cardiovascular effects, the CASAC recommends that additional

studies be included and that the causality determinations be reconsidered in light of these additional studies.

The CASAC commends the thoroughness of the analysis of ecological effects in Appendix 8 and agrees with its causal determinations. However, effects of ozone on wildlife are not characterized. The CASAC recommends that EPA consider developing a research plan for a bird model for toxicology of ozone exposure in warm-blooded vertebrates. Likewise, the CASAC commends the EPA for continuing to clearly characterize and communicate effects of ozone related to climate change in Appendix 9, and agrees with the causal determinations in Appendix 9, but recommends that the EPA consider incorporating further research to better define and quantify the roles of ozone in climate science.

The CASAC appreciates the explanation of the ISA development process in Appendix 10 and notes that parts of it (such as the use of the PECOS tool) appear to be valuable and constitute an advance on earlier approaches. However, as detailed in the consensus responses to the charge questions, it is not clear how, or how well, the approach in Appendix 10 was implemented, and the key conclusions and their rationales in the Draft ISA are unclear to many expert readers. To achieve clearer and more useful and reliable results, the CASAC strongly recommends that the EPA work with the National Academies and external experts in causal analysis, management science, decision analysis, and risk analysis to revise and improve the current ISA development process.

The CASAC appreciates the opportunity to provide advice on the Draft Ozone ISA and looks forward to the agency's response.

Sincerely,

/s/

Dr. Louis Anthony Cox, Jr., Chair
Clean Air Scientific Advisory Committee

Enclosures

NOTICE

This report has been written as part of the activities of the EPA's Clean Air Scientific Advisory Committee (CASAC), a federal advisory committee independently chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. The CASAC provides balanced, expert assessment of scientific matters related to issues and problems facing the agency. This report has not been reviewed for approval by the agency and, hence, the contents of this report do not represent the views and policies of the EPA, nor of other agencies within the Executive Branch of the federal government. In addition, any mention of trade names or commercial products does not constitute a recommendation for use. The CASAC reports are posted on the EPA website at: http://www.epa.gov/casac.

U.S. Environmental Protection Agency
Clean Air Scientific Advisory Committee

CHAIR
Dr. Louis Anthony (Tony) Cox, Jr., President, Cox Associates, Denver, CO

MEMBERS
Dr. James Boylan, Program Manager, Planning & Support Program, Air Protection Branch, Georgia Department of Natural Resources, Atlanta, GA

Dr. Mark W. Frampton, Professor Emeritus of Medicine, Pulmonary and Critical Care, University of Rochester Medical Center, Rochester, NY

Dr. Ronald J. Kendall, Professor of Environmental Toxicology and Head, Wildlife Toxicology Laboratory, Texas Tech University, Lubbock, TX

Dr. Sabine Lange, Toxicology Section Manager, Toxicology Division, Texas Commission on Environmental Quality, Austin, TX

Dr. Corey M. Masuca, Principal Air Pollution Control Engineer, Air and Radiation Protection Program, Environmental Health Services, Jefferson County Department of Health, Birmingham, AL

Dr. Steven C. Packham, Toxicologist, Division of Air Quality, Utah Department of Environmental Quality, Salt Lake City, UT

SCIENCE ADVISORY BOARD STAFF
Mr. Aaron Yeow, Designated Federal Officer, U.S. Environmental Protection Agency, Science Advisory Board, Washington, DC

Consensus Responses to Charge Questions on the EPA's
Integrated Science Assessment for Ozone and Related Photochemical Oxidants
(External Review Draft – September 2019)

Executive Summary

The Executive Summary is intended to provide a concise synopsis of the key findings and conclusions of the Ozone ISA for a broad range of audiences. Please comment on the clarity with which the Executive Summary communicates the key information from the draft ISA. Please provide recommendations on information that should be added or information that should be left for discussion in the Integrated Synthesis and accompanying appendices of the draft ISA.

The key information provided in the EPA's *Integrated Science Assessment for Ozone and Related Photochemical Oxidants (External Review Draft – September 2019)*, hereafter referred to as the Draft Ozone ISA, and its Executive Summary is unclear, for at least the following reasons:

- *Criteria for selecting and weighting studies, and how key conclusions are derived from them, are not clear.* It is unclear how conclusions would change if consistent criteria were systematically applied for selecting, evaluating, summarizing, and synthesizing studies.
- *The Draft Ozone ISA and its Executive Summary do not provide comprehensive quantitative uncertainty and sensitivity analyses* showing how conclusions change for plausible variations in assumptions, interpretations of terms, selection and weighting of studies, and judgments on which the conclusions depend.
- *Causal determination judgments are ambiguous, and sometimes appear subjective and arbitrary*. The Draft Ozone ISA uses the term "causal" and causal determination categories without distinguishing among importantly different causal concepts (e.g., necessary, sufficient, contributing, and other forms of causation). This makes causal statements in the ISA ambiguous, so that it is not possible to determine whether they are correct. Several non-CASAC member consultants commented that different people might make the determinations in very different ways from the same data, and that they could not guess, for any particular body of evidence, which causal determination category EPA would choose to describe it. By this criterion, the causal determinations do not follow clearly from the evidence presented, but incorporate an arbitrary (unpredictable) element.
- *The Draft Ozone ISA's treatment of wildfire contributions to ozone exposure,* and their implications for the National Ambient Air Quality Standards (NAAQS), are unclear.
- *It is unclear to what extent the ozone-associated physiological effects discussed in the Draft Ozone ISA* represent adverse health effects. This is crucial information for policy makers.

The CASAC recommends adding the following information to the Executive Summary:

1. *Discuss how changes in public health effects depend on changes in ozone levels.* This is a crucial scientific topic for informing the Ozone Policy Assessment (PA).
2. *Summarize results from a systematic review, critical evaluation of quality, and synthesis of results from studies informative about public health effects of ozone exposures, including relevant studies that have been omitted from the Draft Ozone ISA* (e.g., Moore et al. 2012). Include results from relevant international and methodological studies (e.g., Vitolo et al., 2018).

Present results of systematic evaluations of individual study quality using consistently applied criteria showing how each individual study is evaluated on each specific quality criterion (e.g., for epidemiological studies, control of observed, residual, and latent confounding; quantification of exposure estimation errors and uncertainties; adjustment of effects estimates for errors-in-variables; quantification of model uncertainty; adjustment of reported significance levels and confidence levels for model uncertainty; control for multiple testing bias; use of appropriate control groups and adjustment sets; tests for internal validity; tests for external validity and invariant causal prediction). The CASAC recommends adding one or more summary displays such as that suggested in public comments from Dr. Julie Goodman (one row per study, one column per criterion, possibly color-coded to show evaluations of quality) to provide insight into the state of the literature and the strengths and limitations of individual key studies.

3. *For epidemiological studies, discuss possible confounding (e.g., by region, season, month, year, day, population, exercise, and other differences) in detail*, and how it was or was not addressed, for individual epidemiological studies and for the epidemiological evidence overall.
4. *Discuss causal biological mechanisms of inflammation-related health effects* and their implications for biologically causal concentration-response (C-R) functions for endpoints, including pulmonary inflammation.
5. *Present results of comprehensive, quantitative uncertainty and sensitivity analyses* showing how conclusions change for variations in inputs, including selection and weighting of studies, alternative interpretations of study results, modeling choices and assumptions, interpretations of terms, and judgments and assumptions on which conclusions depend.

Integrated Synthesis

The Integrated Synthesis presents and synthesizes the overall conclusions from the subsequent detailed appendices of the draft ISA and characterizes available scientific information on policy-relevant issues. Please comment on the usefulness and effectiveness of the summary presentation. Please provide recommendations on approaches that may improve the communication of key findings to varied audiences and the synthesis of available information across subject areas. What information should be added or is more appropriate to leave for discussion in the subsequent detailed appendices?

The Integrated Synthesis has the following limitations that reduce its usefulness and effectiveness:

- *Omission of studies.* Multiple readers of the Draft Ozone ISA, including several non-CASAC member consultants, noted that it omits many relevant studies, especially many that do not conclude that ambient ozone causes adverse health effects.
- *Literature on nonlinear effects is not well covered.* The Draft Ozone ISA does not adequately cover the recent scientific literature on nonlinear C-R functions for ozone. For example, p. IS-88 states that "Across recent studies that used a variety of statistical methods to examine potential deviations from linearity, **evidence continues to support a linear C-R relationship**, but with less certainty in the shape of the curve at lower concentrations (i.e., below 30−40 ppb)." This contrasts with a substantial international literature, disregarded in the Draft Ozone ISA, on nonlinear C-R relationships (e.g., Bae et al., 2015; Seltzer et al., 2018; Wilson et al., 2014).
- *Summaries of relevant literature are incomplete and of questionable accuracy.* The Draft Ozone ISA does not provide a complete and accurate summary of the studies that it cites. For example:

- Page 3-91 of the Draft Ozone ISA states that "A recent [Children's Health Study] CHS analysis examined asthma incidence in relation to improved air quality in nine southern California communities (Garcia et al., 2019). ***Decreases in baseline ozone concentrations in three CHS cohorts, enrolled in 1993, 1996, and 2006, were associated with decreased asthma incidence.***" However, Garcia et al. (2019) state that "Among children in Southern California, decreases in ambient nitrogen dioxide and $PM_{2.5}$ between 1993 and 2014 were significantly associated with lower asthma incidence. ***There were no statistically significant associations for ozone*** or PM_{10}." (Garcia et al., 2019, emphasis added.)
- Table 3-3 on "Summary of evidence for a likely to be causal relationship between long-term ozone exposure and respiratory effects" cites the study of Moore et al. (2008) as providing "key evidence" for the ISA's causal determination that there is "a likely to be causal relationship between long-term ozone exposure and respiratory effects." Specifically, Moore et al. is cited as providing "***Consistent evidence of an association between long-term ozone concentrations and hospital admissions*** and [Emergency Department] ED visits for asthma." Yet, follow-up work by Moore et al. (2012) noted methodological limitations of the 2008 paper (especially, that its results may have resulted from incorrect untested modeling assumptions, rather than from information in the data) and provided and applied an improved methodology ("CMRIER" or "causal models for realistic individualized exposure rules"). A key result was that the previous significant effect of ozone was no longer found (Moore et al., 2012). However, this may have been due to the reduced power from reducing the dataset. This more recent paper is not mentioned in the Draft Ozone ISA. The Draft Ozone ISA cites the 2008 results as "key evidence" without noting that the authors subsequently arrived at a different conclusion in the 2012 paper.
- Table 3-3 cites a study by Tétreault et al. (2016) as providing "Key Evidence" of "Cohort studies demonstrating an *association* with asthma development in children." The Draft Ozone ISA then interprets this, without any detailed explanation, as "Evidence for a *likely to be causal* relationship between long-term ozone exposure and respiratory effects." (Emphases added.) But it is not clear how or whether the Draft Ozone ISA considered the results of sensitivity analyses for the individual studies it relies on for its conclusions, in interpreting the Tétreault et al. (2016) study as "Key Evidence" of a "likely to be causal" relationship; or how sensitive the resulting causal determinations are to incompletely controlled confounding.

- *The science related to possible health benefits of reducing ozone needs to be more fully addressed.* The Draft Ozone ISA does not usefully summarize, or critically evaluate, available scientific information on whether or to what extent reducing ozone reduces public health risks. Yet, this is a crucial topic needed to inform policy decisions about the public health consequences of alternative possible policy choices. For example, the non-CASAC member consultants were directly asked "*Can valid determinations of manipulative or interventional causation – that is, how and whether changing exposure would change health risks – be made based on observed associations of the types analyzed in the ISA*?" Most who answered said no; none said yes (see responses in Appendix B). Unless this omission is fixed, the PA lacks a scientific foundation in the ISA for predicting effects on public health of alternative policies.

As mentioned above, the following additions to the Draft Ozone ISA and Executive Summary are recommended to improve the communication of key results, and also the policy relevance, scientific validity, and methodological integrity of the content being communicated:

1. *Summarize available empirical evidence on how changes in public health effects depend on changes in ozone levels.*
2. *Present summary results from a systematic review and critical evaluation and synthesis of relevant studies*, including negative ones that have been omitted from the Draft Ozone ISA.
3. *Provide detailed discussion of possible confounding*, and how it was or was not addressed for each study used to support causal conclusions.
4. *Present results of systematic evaluations of study quality*, using consistently applied criteria, showing how each key study included performs on each specific quality criterion relevant for drawing valid causal conclusions.
5. *Discuss causal biological mechanisms of inflammation-related health effects* preventable by reducing current ozone levels.
6. *Present comprehensive, quantitative uncertainty and sensitivity analyses* showing how the ISA's conclusions change for variations in selection and weighting of studies, modeling choices and assumptions, interpretations of undefined and vague terms, and subjective judgments on which the conclusions depend.

Appendix 1

To what extent is the information presented in Appendix 1 regarding sources, precursor emissions, and measurement and modeling of ambient concentrations, as well as modeled estimates of background concentrations of ozone, clearly and accurately conveyed and appropriately characterized? Please comment on the extent to which available information on the spatial and temporal trends of ozone concentrations at various scales has been adequately and accurately described.

Section 1.3 (Sources of U.S. Ozone and its Precursors) presents estimated national values for 2014/2017 National Emissions Inventory (NEI) emissions. However, there is no detailed discussion on the uncertainty associated with each pollutant or source sector. Some pollutants and sectors will be much more uncertain than others. For example, oxides of nitrogen (NOx) emissions from electric generating units (EGUs) have low uncertainty since they are typically measured by hourly continuous emissions modeling (CEMs). On the other hand, other source sectors and pollutants may be highly uncertain. The uncertainties in the emissions inventory (magnitude, spatial allocation, and temporal allocation) should be discussed for each pollutant and source sector. In addition, it would be helpful to add national maps containing county-level emissions for NOx, volatile organic compounds (VOCs), carbon monoxide (CO), and methane (CH_4) to show the variability across the country.

It is not clear if CH_4 is included in the VOC emissions or not. The text should clearly state if CH_4 is included or excluded from the VOC emissions discussed in this appendix. Due to the importance of biogenic VOCs, this section should discuss the differences between the Biogenic Emission Inventory System (BEIS) and Model of Emissions of Gases and Aerosols from Nature (MEGAN) models that are typically used to estimate biogenic VOC emissions. In addition, biogenic VOC trends should be included to see the variability from year-to-year and season-to-season.

Section 1.4 (Ozone Photochemistry) should start with a discussion of why the precursor emissions discussed in Section 1.3 (NOx, VOCs, CO, and CH_4) are important for ozone formation. An overview of the chemical mechanism should be presented, and important chemical reactions should be highlighted. The relative importance of each ozone precursor should be discussed relative to local ozone formation (both urban and rural) in comparison to U.S. background ozone formation.

Section 1.5 (Inter-Annual Variability and Longer Term Trends in Meteorological Effects on Anthropogenic and U.S. Background Ozone) should discuss the impact of inter-annual variability and longer term trends in meteorological effects on ozone design values. In addition, this section should add a discussion on the topographical effects on meteorology, ozone formation, and ozone transport.

Section 1.6 (Measurements and Modeling) should discuss ground-based ozone lidar instruments that measure the vertical structure of ozone and quantify the mixing of plumes aloft. A review of these instruments and their capability should be added to this section. The section on "Satellite-Based Remote Sensing Methods" should include a discussion of the new TROPOspheric Monitoring Instrument (TROPOMI) satellite data that includes high resolution measurements of nitrogen dioxide (NO_2) and formaldehyde. The section on "Advances in Regional Chemical Transport Modeling" should discuss the importance of performing a comprehensive model performance evaluation when using regional chemical transport models. This evaluation should include an evaluation of precursor pollutants to help ensure the model does not have compensating errors.

EPA's 2016 Exceptional Events Rule allows certain ozone measurements due to natural events to be excluded from the official design values when compared to the NAAQS. In some cases, identical exceptional events can be treated differently in one location vs. another based on how close the area is to the standard. In both locations, people could potentially be impacted by adverse health effects from ozone, but the data is removed in one location and not the other. The Draft Ozone ISA should discuss how exceptional events are accounted for in health studies and risk analyses.

Section 1.7 (Ambient Air Concentrations and Trends) should discuss the shifting of ozone peak concentrations from summer to spring and fall that is occurring in many parts of the country (Blanchard and Hidy, 2018; Blanchard et al., 2019). In addition, this section should include a discussion on ozone precursor trends in addition to ozone trends. Specifically, trends in NOx, VOCs, and CO measurements form national monitoring networks (AQS, near-road, NCore, and PAMS) should be included and discussed.

Section 1.8.1 (Modeling Strategies Applied to Estimate U.S. Background Ozone) begins with the statement "As described in Section 1.2.2.1, [U.S. Background] USB ozone cannot be reliably estimated using ambient monitoring data because monitors can be influenced by U.S. emissions, including both relatively nearby emissions and interstate and hemispheric transport of ozone produced from U.S. emissions." Parrish et al. (2017) and Parrish and Ennis (2019) have shown that USB ozone can be reliably estimated using ambient monitoring data. Although monitors can be influenced by U.S. emissions, it is possible to account for these influences. Estimates from measurement-based approaches and from modeling-based approaches can be compared to understand differences and minimize the uncertainty in USB ozone estimates.

Emission controls have reduced ozone in the United States to the extent that background ozone contributes the majority of urban ozone concentrations, even on many days when ozone exceeds the NAAQS. Figures 1 and 2 show estimates of the ozone design values that would be present in the absence of U.S. or North American anthropogenic emissions. Figure 1 is from a model calculation using the "zero-out sensitivity approach" (Jaffe et al., 2018). Figure 2 is developed from an observational-based approach (Parrish et al., 2017; Parrish and Ennis, 2019) applied to the entire country. These two maps show that in the southwestern United States, background ozone levels are relatively high, close to 70 ppb. Section 1.8.2.1 discusses new USB and North American Background estimates, but all of these estimates are for seasonal means. The EPA should also discuss the ozone design values that can result from USB.

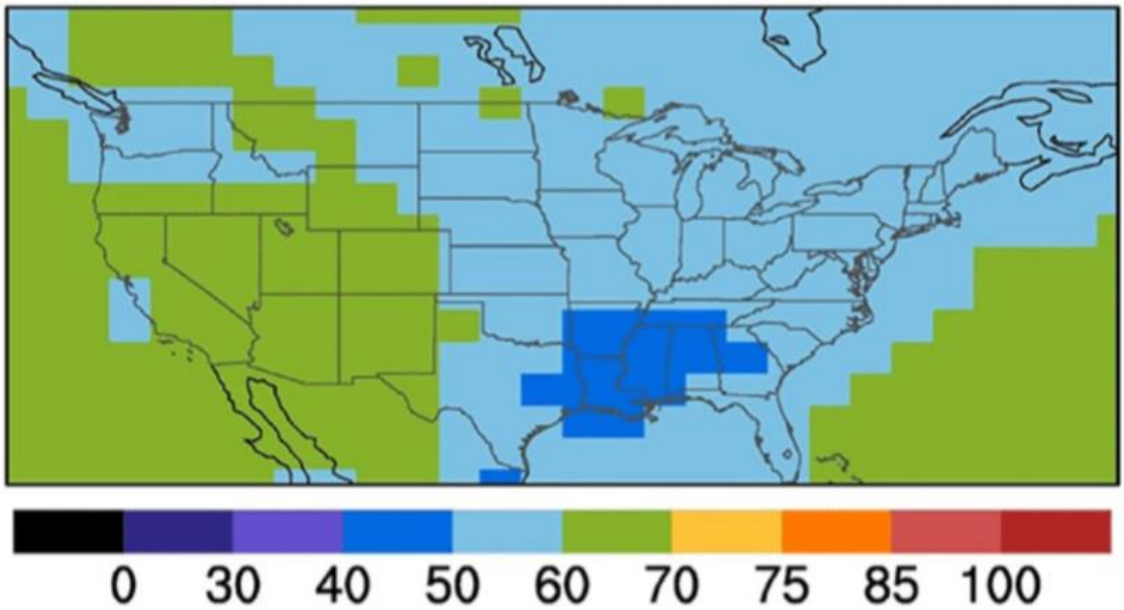


Figure 1. Annual 4th highest MDA8 O3 in ppb from North American background (i.e., with North American anthropogenic precursor emissions set to zero) averaged over 2010–2014 from a GFDL-AM3 model simulation (Jaffe et al., 2018).

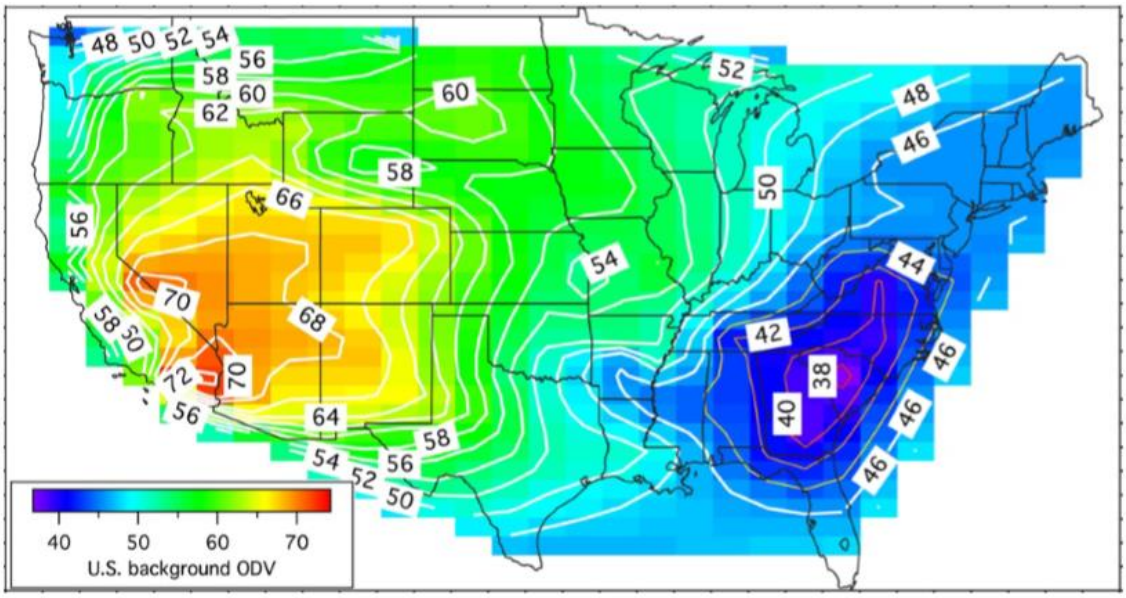


Figure 2. Ozone design values expected from U.S. background (i.e., with U.S. anthropogenic precursor emissions set to zero) in ~ 2015 derived from observations (D.D. Parrish, unpublished figure).

Appendix 2

Appendix 2 describes scientific information on exposure to ozone and implications for epidemiologic studies. To what extent is the discussion on methodological considerations for exposure measurement and modeling clearly and accurately conveyed and appropriately characterized? Please comment on the extent to which the discussion regarding exposure assessment and the influence of exposure error on effect estimates in epidemiologic studies of the health effects of ozone has been adequately and accurately described.

Section 2.3 (Exposure Assessment Methods) gives a high-level overview of fixed-site monitors, passive and active personal samplers, spatial interpolation, land use regression and spatiotemporal modeling, chemical transport modeling, hybrid approaches, and microenvironmental modeling. The discussion on microenvironmental modeling should include additional information on the Air Pollution Exposure model (APEX) and Stochastic Human Exposure and Dose Simulation (SHEDS) models.

Section 2.4 (Personal Exposure) discusses updates to the Consolidated Human Activity Database (CHAD), indoor-outdoor (I/O) ratios to describe infiltration of ambient ozone into homes and buildings, and personal exposure-ambient concentration (P/A) ratios where an individual is exposed. Additional discussion should be added for ozone infiltration in vehicles since a large amount of time is spent commuting. Also, a detailed discussion of the uncertainties and variability associated with the CHAD, I/O ratios, and P/A ratios should be included.

It is stated in Section 2.5 (Copollutant Correlations and Potential for Confounding):

> "Given that the majority of the copollutant correlation data are low, confounding of the relationship between ambient ozone exposure and a health effect by exposure to CO, SO_2, NO_2, PM_{10}, or $PM_{2.5}$ is less of a concern for studies of the health effects of ambient ozone exposure

compared with studies of the health effects related to exposure of other criteria air pollutants. When copollutant correlations are higher during the warm season, greater risk of copollutant confounding exists."

However, the summer is the season with the highest ozone concentrations and the highest ozone exposure; therefore, a greater risk of copollutant confounding exists and should be accounted for in the interpretation of the epidemiological results.

The final ISA should include a more detailed discussion of the impact of exposure measurement error on effect estimates in epidemiology studies. Estimation errors typically lead to overestimates of low-dose risks and underestimates of high-dose risks if the true causal C-R function has a threshold or threshold-like nonlinearity. Many studies have shown that bias or error in the exposure or outcome assignment can cause the estimated C-R function to flatten and appear linear even if the true C-R function has a well-defined threshold or other non-linear shape (Brauer et al., 2002; Cox, 2018; Lipfert and Wyzga, 1996; Rhomberg et al., 2011; Watt et al., 1995; Yoshimura, 1990). Studies have also shown that exposure error can in fact have complicated effects on a health effect estimate that are not captured by the generalization that the effect is underestimated (Cefalu and Dominici, 2014; Goldman et al., 2011; Jurek et al., 2008, 2005; McGuinn et al., 2017; Sorahan and Gilthorpe, 1994). Studies that have investigated the effects of better exposure estimates on health effect estimates (e.g., Ebelt et al., 2005; Hart et al., 2015; McGuinn et al., 2017; Trenga et al., 2006) have demonstrated that there is no or little difference in health effects estimates or width of confidence intervals with different (presumably better) exposure estimates. In copollutant models, whichever pollutant is measured with the least error is most likely to be ascribed the positive effect. This phenomenon has been demonstrated by several groups (Carrothers and Evans, 2000; Fewell et al., 2007; Lipfert and Wyzga, 1996), is discussed in the Draft Particulate Matter Integrated Science Assessment (US EPA, 2018), and it makes interpreting copollutant models quite challenging. Addressing this problem requires considerations of joint exposure measurement errors for each component.

The summary table provided in the EPA Ozone ISA presentation to CASAC on December 4, 2019, showing the influence of exposure error on epidemiology study outcomes (page 18) is a very useful summary and should be included in Section 2.6 (Interpreting Exposure Measurement Error for Use in Epidemiology Studies).

Appendices 3-7

Please comment on the identification, evaluation and characterization of the available scientific evidence from epidemiologic, controlled human exposure, toxicological and associated human exposure and atmospheric sciences studies and the application of information from these studies to inform causality determinations for human health outcomes.

Appendices 3-7 present assessments of the health effects associated with short-term and long-term exposure to ozone. The discussion is organized by exposure duration, broad health effects (e.g., asthma, ischemic heart disease, etc.), and scientific discipline. Please comment on the characterization of the evidence within these chapters.

Please comment on the portrayal and discussion of the biological plausibility evidence presented in Appendices 3-7 and the extent to which: (1) the organization adequately captures the current state of the science with respect to potential pathways by which ozone could impart health effects, and (2) as currently constructed, inform causality determinations.

Study Quality

The CASAC recommends that the EPA explicitly address chance, bias, confounding, and other non-causal sources of associations (e.g., historical trends, model misspecification, measurement error, etc.) in their study quality analysis of epidemiology studies in the Draft Ozone ISA. These (in addition to causality) are all potential reasons for an epidemiology study to observe an association between two variables (Zaccai, 2004) and therefore should be more explicitly considered when presenting and discussing study results. In addition, factors other than just copollutants should be considered as important confounders in the referenced epidemiology studies. To properly consider chance, results that are not statistically significant should be indicated as such when results are discussed. If there is a reason why statistical significance may not have been achieved (e.g., low sample size), this should be included in the discussion of the study results. The general conclusion from the non-CASAC member consultants was that statistical significance does need to be given some consideration, in addition to other factors such as patterns in the epidemiology data.

Accuracy of Presentation

The EPA should provide a balanced summary of the study results for each health endpoint. Adequately communicating available positive, negative, and null results provides useful information for further documents in the Ozone NAAQS review.

In section summaries, divergent results should not be ignored, but rather should be included in a more nuanced summary of results. For example, the Arjomandi et al. (2018) study did not find an association between GSTM1 genotype and ozone-induced airway inflammation. However, in the summary section for respiratory effects in healthy populations this divergent finding was not included: The Draft Ozone ISA states that "Recent studies are consistent with previous findings and expand on observed interindividual variability in inflammatory responses, providing additional evidence that GSTM1-null individuals are more susceptible to ozone-related inflammatory responses." This statement is not an accurate summary of the newly available results and should acknowledge the findings of Arjomandi et al. (2018). Similarly, page 3-91 of the Draft Ozone ISA states that "A limited number of recent studies provide evidence of an association between long-term exposure to ozone and asthma development in children. … An overview of the evidence is provided below. A recent CHS analysis examined asthma incidence in relation to improved air quality in nine southern California communities (Garcia et al., 2019). **Decreases in baseline ozone concentrations in three CHS cohorts, enrolled in 1993, 1996, and 2006, were associated with decreased asthma incidence.**" (Emphasis added.) However, Garcia et al. (2019) state that "Among children in Southern California, decreases in ambient nitrogen dioxide and $PM_{2.5}$ between 1993 and 2014 were significantly associated with lower asthma incidence. **There were no statistically significant associations for ozone** or PM_{10}." (Emphasis added.)

Further, information summarized from one section to another should maintain the accuracy and nuance of the underlying data. For example, in Sections 4.1.16 and 6.2.4.1, the EPA states "Specifically, the evidence from controlled human exposure studies provided support for increased decrements in FEV_1

and greater inflammatory responses to ozone in individuals with asthma than in healthy individuals without a history of asthma." Although some studies have found that people with asthma have greater lung function decrements in response to ozone, compared to people without asthma, other studies have noted no difference in the two populations in response to ozone. The respiratory chapter addresses this point at length, and states that people with asthma are at least as sensitive as people without asthma to lung function effects of ozone. Similarly, in the summary and causality determination section for long-term total mortality, the following statement contrasts with previous text and the overall conclusions: "There is coherence across the scientific disciplines (i.e., animal toxicology, controlled human exposure studies, and epidemiology) and biological plausibility for ozone-related cardiovascular (Appendix 4) and respiratory (Appendix 3) endpoints, which lend some additional support to the ozone-mortality relationship." The point is made repeatedly earlier in the Draft Ozone ISA that the clinical studies are inconsistent with regard to cardiovascular effects. This sentence needs to be reconsidered and harmonized with the rest of the document.

The CASAC recommends that the EPA carefully review the Draft Ozone ISA for accuracy. The EPA should provide accurate study information as well as study conclusions that are consistent with the conclusions made by the study authors (and if they are not consistent, the EPA should explain why they have a different interpretation of the results than the study authors). One method for ensuring that the summarized information is consistently accurate is to use random spot-checking for data accuracy, to ensure that summaries reflect the study conclusions, and to establish consistency between chapters. The study results presented in the metabolic chapter are particularly error-prone. For example,

- In Section 5.1.4 (overweight and obesity) the EPA provides incorrect exposure information for the Gordon et al. (2016) study (animals exposed one day per week, not 4 days per week).
- In Section 5.1.5.1 (Other indicators of metabolic function, inflammation), the EPA states that "Obesity-prone mice (adult male KK mice) were exposed to ozone for 13 consecutive weekdays [4 hours/day; Zhong et al. (2016)]." However, Tables 5-7 and 5-10 indicate that the exposure was 3 consecutive days. The exposure regimen in Zhong et al. (2016) was 13 consecutive weekdays.
- In Section 5.1.5.4 (Other indicators of metabolic function, serum lipids), the information presented for the Gordon et al. (2016) study is inaccurate. The EPA states that "The effect of high-fat and high-fructose diets was tested in male brown Norway rats" – the study was done in male and female rats; "With ozone exposure (0.8 ppm ozone, 4 days/week for 3 weeks)," – exposure was 0.8 ppm ozone, 5 hrs/day, 1 day/week for 4 weeks (subacute exposure), or a single 0.8 ppm exposure for 5 hrs (acute exposure); "Females were refractory to change" – the abstract of the paper says "Female rats appeared to be more affected than males to O3 regardless of diet."
- In Section 5.2.5 (metabolic syndrome and type II diabetes), the Jerrett et al. (2017) effect estimates are incorrect (presented are 1.28; 95% CI: 1.06, 1.55 and 1.20; 95% CI: 0.96, 1.50 with NO_2 adjustment – should be 1.18 (1.04, 1.34) and 1.13 (0.97, 1.31) with NO_2 adjustment).

Dose Assessment and Concordance

The EPA should appropriately compare animal to human ozone doses when extrapolating animal exposures to potential human risks. The Population, Exposure, Comparison, Outcome and Study Design (PECOS) statement for experimental studies in Appendix 3 (and on pp. 3-19 and 3-26) notes that resting rats exposed to 2 ppm have an equivalent ozone deposition as exercising humans, citing Hatch et al. (1994). The EPA should further discuss that there is a similar alveolar dose of ozone at equal ozone

exposure concentrations when humans and rats are both at rest, and that a human with a ventilation rate that is five times higher than resting will have a 5 times higher dose. This should be correctly noted and Hatch et al. (2013) and McCant et al. (2017) (which describes this misconception) should be cited.

The EPA should also present dose information in their biological plausibility discussion because of concerns about dose-dependent transitions in toxicity: that is, the principle that mechanisms of toxicity can change with different doses of a toxicant (Slikker et al., 2004a, 2004b). By incorporating dose (or concentration and time) information into the biological plausibility discussions, the EPA will allow the reader to judge whether there is uncertainty in the application of a particular biological mechanism to an ambient human exposure. The non-CASAC member consultants also generally agreed that, given that the causality determination for metabolic effects of ozone exposure is mostly derived from animal toxicological studies, it is appropriate for the EPA to more thoroughly discuss the dosimetric similarities and differences between animals and humans, beyond simply referencing Hatch et al. (1994).

Clarity of Presentation

The EPA should clearly present study information, results, and discussion in each of the Draft Ozone ISA sections, and should provide an accurate and balanced summary of results. When discussing the results from all studies, and particularly controlled human exposure (CHE) studies, it is important to include the exposure duration (e.g., on p. 3-26 when discussing concentrations at which airway hyper-responsiveness has been observed) and the exercise level of the participants (e.g., in the integrated synthesis when discussing concentrations that could generate adverse effects in healthy adults).

For the discussions addressing pre-existing conditions, the EPA should specifically include and address data that provides information on responsiveness of people with the condition compared to people without the condition (because this directly informs potentially sensitive subpopulations). For example,

- In Section 3.1.6.2 the EPA addresses respiratory effects of ozone in people who are obese or who have metabolic syndrome. However, in describing the study results, particularly of Ying et al. (2016), Zhong et al. (2016), or Gordon et al. (2016b), the EPA does not note whether there was a greater (or different) inflammatory response to ozone in the obese/metabolic syndrome animals versus lean/healthy animals. Because this distinction seems to be the purpose of this section, these pieces of information should be included.
- The EPA includes sections about respiratory effects of ozone with exposure during pregnancy (3.2.4.7) and in populations with metabolic syndrome (3.2.4.8). Is the purpose of these sections to show that there is an increased response to ozone in these populations? If so, then the EPA should specifically provide information and discuss whether the data show that these groups are more sensitive. As it stands, this conclusion is not clear.
- In Section 5.1.4 (overweight and obesity), there is a mixing of concepts that is confusing and perhaps misleading. Some of the studies summarized here are relevant to obesity as a risk factor, in other words, whether obesity as a subject characteristic enhances ozone responses: pulmonary, cardiovascular (CV), or other. Descriptions of these studies should be part of the other chapters in the sections discussing sensitive subpopulations with regard to those outcomes. The issue being considered in this section is whether ozone alters metabolic functions including body weight, body mass index (BMI), body composition, caloric intake, glucose metabolism, lipid metabolism, stress responses, etc. The sentence in this paragraph starting on page 5-12, line 12, describes what this section should be about: “Recent toxicological studies provided some

evidence that ozone may impair metabolism and affect body weight, BMI, and body composition, as well as effect [sic] caloric intake."

If possible, the EPA should avoid making statements that address an unlikely conclusion, but that avoid addressing the conclusion of interest. For example, in Section 3.1.5.4 (lung function): "it was concluded that individuals with asthma were *at least as sensitive* to acute effects of ozone as healthy individuals." (Emphasis added). The conclusion of interest is whether people with asthma are *more* sensitive or not.

The EPA's underlying concern about people with asthma is not only that they will have an increased innate response to ozone exposure (although the data are unclear about whether effects on lung function, airway inflammation, or airway responsiveness are greater in asthmatic compared with healthy subjects), but rather that people with asthma likely have less of a buffer against adverse effects. This important argument deserves more emphasis when discussing the respiratory effects of ozone exposure on people with asthma.

The EPA has described the exercise level in key CHE studies such as Schelegle et al. (2009) as a slow walking pace, but Schelegle et al. (2009) write that "This protocol contains six 50-minute exercise periods with minute ventilation maintained at 8 L/min/L of FVC (VE of approximately 40 L/min). As noted by Folinsbee et al. (1988) and McDonnell et al. (1991), this level of exertion was ''intended to simulate work performed during a day of heavy to severe manual labor in outdoor laborers.'" This discrepancy in description of the exercise level should be clarified.

In Section 3.2.1 (long term ozone exposure and respiratory effects), the first paragraph includes a summary of the findings from the 2013 Ozone ISA. This summary should include the limitations and uncertainties which at that time precluded a determination of "causal" for long-term respiratory effects.

In Section 4.1.8 (blood pressure changes and hypertension), when discussing emergency department (ED) visits and hospital admissions (HAs), the EPA describes the study findings in the context of the mean ozone concentrations measured in the study areas. Including mean measured concentrations in the discussion of study results is very helpful and would be a valuable addition to other sections of the Draft Ozone ISA.

In Section 4.1.9.2 (heart rate and heart rate variability), the EPA describes results from the Arjomandi et al. (2015) study as showing *associations* between exposure and measured effects. However, Arjomandi et al. (2015) is a controlled and blinded experimental human exposure study where the subjects were exposed to clean air and 2 concentrations of ozone for 4 hours, with intermittent exercise, with heart rate variability measured before and at intervals after exposure. This paragraph should be re-written to indicate that the changes can reasonably be described as *effects* of the exposure, rather than *associations*.

When the EPA states that there is little evidence for ozone impacting a particular endpoint, they should clarify whether there is little evidence because studies have not been done to investigate ozone effects on that endpoint, or if the available studies do not show an association or an effect of ozone exposure. For example, in Section 4.2.2 (long-term cardiovascular effects, biological plausibility) the EPA notes that "However, considerable uncertainty remains in how long-term ozone exposure may lead to mortality given that there is little epidemiologic evidence of an association between long-term exposure to ozone and other cardiovascular endpoints such as IHD, stroke, or thromboembolic disease." This statement

should be clarified to specify whether the studies have not been done, or if studies have been done but have not shown associations.

Consistency of Results & Reporting

The Draft Ozone ISA would be strengthened if the EPA discussed the scientific significance of conflicting and/or inconsistent evidence. An example of inconsistent (or seemingly inconsistent) results comes from Section 5.2.3 (long-term exposure, glucose and insulin homeostasis) where evidence is presented from three studies (Miller et al., 2016b; Gordon et al., 2013; and Bass et al., 2013). These studies tested effects of long-term ozone exposure in male rats. But they show different effects: Bass et al. (2013) showed no change in fasting glucose with subchronic exposure, but Miller et al. (2016b) did; Miller et al. (2016b) showed decreased baseline insulin in subchronically-exposed adult animals, but Gordon et al. (2013) showed no change in adult exposed animals, and increases in insulin in senescent exposed animals. The EPA could speak to whether there are patterns in these results, or if the differences are spurious or related to strain differences.

Applicability of Results from Animal Studies

Dose-responsiveness of effects of ozone exposure in experimental studies can be used to identify relevant biological plausibility pathways and exposure-specific responses, and so should be further discussed in those sections. In particular, establishing no-effect and low-effect concentrations for endpoints such as long-term ozone exposure and lung function development would ease the extrapolation of results from animal toxicology experiments to effects in humans at ambient concentrations.

In addition, information about the comparability of animal models to human disease are useful in extrapolating results from animal studies – such as information about how the mouse model of allergic airway disease compares to asthma in humans. Also important is information allowing the interpretation of *ex vivo* studies, such as experiments in isolated, perfused hearts (Section 4.1.4, heart failure, impaired heart function, and associated cardiovascular effects).

Shape of the C-R Function

As was discussed in the CASAC's review of the Particulate Matter ISA and PA, errors and heterogeneity in epidemiology study variables can affect the apparent shape of C-R relationships and can obscure thresholds. Evidence for this has been provided by many peer-reviewed publications (Brauer et al., 2002; Cox, 2018; Lipfert and Wyzga, 1996; Rhomberg et al., 2011; Watt et al., 1995; Yoshimura, 1990) and notably by the EPA in the ISA preamble (U.S. EPA 2015, Section 6c, p. 29):

> "Various sources of variability and uncertainty, such as low data density in the lower concentration range, possible influence of exposure measurement error, and variability among individuals with respect to air pollution health effects, tend to smooth and 'linearize' the concentration-response function and thus can obscure the existence of a threshold or nonlinear relationship. Because individual thresholds vary from person-to-person due to individual differences such as genetic differences or pre-existing disease conditions (and even can vary from one time to another for a given person), it can be difficult to demonstrate that a threshold exists in a population study. These sources of

> variability and uncertainty may explain why the available human data at ambient concentrations for some environmental pollutants (e.g., PM, O_3, Pb, environmental tobacco smoke, radiation) do not exhibit population-level thresholds for cancer or noncancer health effects, even though likely mechanisms include nonlinear processes for some key events."

The problem described here is not whether a threshold in the data may exist, but rather that even if it does exist, epidemiology studies may not be capable of definitively identifying the threshold. To address this concern the EPA should explicitly acknowledge in the Draft Ozone ISA that variability and error in the variables can linearize C-R functions and obscure thresholds, and this acknowledgement should be included in those places where the EPA concludes that the relationship between ozone and a health effect is linear and has no threshold. The CASAC also recommends that the EPA begin to apply methods (and encourage the epidemiological community to apply methods) to address this particular concern, including errors-in-variables methods. If possible, the EPA should include these types of adjustments when applying epidemiology C-R functions to their risk assessments.

In Section 6.1.7 (short-term ozone exposure and mortality, shape of the C-R function), the EPA states that in the previous ISA the available studies showed no evidence of a deviation from linearity or the presence of a threshold for short-term ozone-mortality relationships. "However, it is important to note that the examination of the ozone-mortality C-R relationship is complicated by previously identified city-to-city and regional heterogeneity in ozone-mortality risk estimates (U.S. EPA, 2013). Recent studies continue to provide evidence of a linear C-R relationship with no evidence of a threshold below which mortality effects do not occur along the distribution of ozone concentrations observed within the U.S." The EPA should provide information in this section noting whether the new studies address the considerations of city-to-city or regional heterogeneity that were concerns before, or if this is still an issue. If it is, the EPA should state it as such.

In addition, some of the plots that are presented by the EPA do not visually appear to be linear and do appear to have a threshold, such as the Silverman and Ito (2010) plot (Figure 3-9), the Moolgavkar et al. (2013) plot (Figure 6-6) and the Di et al. (2017) plot (Figure 6-7). If the EPA thinks that there is so much uncertainty at the lower ends of these curves that we cannot trust the apparent non-linear shape, then we also cannot trust that the shape is linear, and no conclusions should be drawn about shape at low concentrations.

In addition, the Draft Ozone ISA does not adequately cover the recent scientific literature on nonlinear C-R functions for ozone. This includes work by: Bae et al. (2015) who report that "The mean O_3 concentration did not differ greatly between Korea and Japan, which were 26.2 ppb and 24.2 ppb, respectively. Seven out of 13 cities showed better fits for the spline model compared with the linear model, supporting a **non-linear relationships between O_3 concentration and mortality**. **All of the 7 cities showed J or U shaped associations suggesting the existence of thresholds**. The range of city-specific thresholds was from 11 to 34 ppb. **The city-combined analysis also showed a non-linear association with a threshold around 30-40 ppb**." (Emphasis added); and Wilson et al. (2014) who report that, even in modeling that constrains ozone C-R functions for mortality to be monotonic (disallowing J-shaped or U-shaped relations such as those reported by Bae et al., 2015), "We then examine the synergistic effect of ozone and temperature both nationally and locally and **find evidence of a nonlinear ozone effect** and an ozone-temperature interaction at higher temperatures and ozone

concentrations." (Emphasis added.) The draft ISA reports the nonlinear interaction from this study (p. 6-12) but does not mention the "evidence of a nonlinear ozone effect."

Interpretation of Study Results

The Draft Ozone ISA should address the adversity and clinical significance of important health effects, such as changes in fasting blood glucose (presented in Section 5.1.3 - glucose and insulin homeostasis). In addition, inclusion of the significance of relationships between different factors identified in key studies would help clarify the conclusions that can be drawn. For example, the significance of the relationship between new-onset asthma and ozone in children with various genetic variants (Section 3.2.4.1), the association between changes in heart rate or blood pressure and ozone in people with mood disorders (Section 4.1.16), or the increase of inflammatory mediators in epididymal adipose (Section 5.1.5.1).

Completeness of Study Information

The EPA should ensure that all relevant information is included in the study figures or tables. For example, in Section 3.1.10.1 (short-term respiratory effects, copollutant confounding), the EPA notes that they provide study-specific details in the tables in Section 3.3. However, the information in those tables do not include the effect estimates for the copollutant models, only for the single pollutant models. The EPA should include the copollutant effect estimates in these tables, or in the text or figures of this section. The latter would be preferable, because of the importance of considering copollutant confounding. Similarly for the results that consider confounding by aeroallergens. In addition, Table 5-7 does not include all of the information about the Ramot et al. (2015) study – only one rat strain is included and not the eight that were tested, and only one of the three ozone doses is included.

Causality Determinations

For the short-term ozone effects on metabolic endpoints, the causality determination of "likely to be causal" is not warranted. The studies often do not find consistent direction of effects on the measured endpoints – if biomarkers change in different directions in different experiments, does that matter for the EPA's causality determination? In general, this causality determination is driven by the animal toxicology, which is largely limited to rodents. The animal data on glucose and insulin effects seem to be fairly robust, but the extrapolation of the findings to humans is in question. The epidemiological evidence is sparse and inconsistent, without any evidence of adverse clinical outcomes related to metabolic effects. The only human clinical study (Miller et al., 2016a) showed no effects on insulin levels or homeostatic model assessment for insulin resistance (HOMA-IR), but did find acute increases in stress hormones in response to ozone exposure. It is as yet unconfirmed. While the animal studies provide plausibility, the sparse epidemiology and human clinical data do not justify the "likely" determination. "Suggestive" appears to be more appropriate.

Overall for the causality designation for long-term ozone effects on metabolic endpoints, there is limited epidemiology evidence, and those data are not clear-cut. For example, some associations are lost with copollutants added to the models, or copollutants are not assessed, study quality was only assessed in one of the six epidemiology studies cited in Table 5-4, and at least one of the study quality details for that study in the Health Assessment Workspace Collaborative (HAWC) was incorrect. The animal evidence is not always summarized correctly and shows somewhat inconsistent results. It does

consistently show no effects at lower ozone concentrations (0.25 ppm), and all 3 of the cited animal studies were conducted in whole or in part by the same group of authors. As with the short-term designation, the evidence does not justify the "likely" determination.

For ozone and effects on fertility and reproduction, the effects of ozone on male reproduction are based on little data (inconsistent epidemiology studies, one animal study), and the EPA states for female reproduction that "In conclusion, results from epidemiologic studies are mixed, with benefits and detriments to female reproductive function with ozone exposures, while toxicological studies show limited evidence of effects on successful completion of pregnancy." Therefore, it is not clear why the EPA has designated fertility and reproduction as "suggestive of causality." The CASAC suggests that this causality designation should be reconsidered by the EPA.

Study Inclusion

Section 10.3.1.4 (study selection: full-text evaluation of studies, relevance) indicates, "In instances when a "causal" or "likely to be a causal" relationship was concluded in the 2013 Ozone ISA (i.e., short-term ozone exposure and respiratory and cardiovascular effects and total mortality, and long-term ozone exposure and respiratory effects), the epidemiologic studies evaluated for those outcomes were more limited in scope and targeted towards study locations that include U.S. airsheds or airsheds that are similar to those found in the U.S., as reflected in the PECOS tool." The rationale for limiting epidemiology studies in these categories of causality is to emphasize the studies most relevant for policy in addressing possible changes in the NAAQS. This is reasonable for outcomes determined to be *causal* or *likely to be causal*. The problem is that, in the Draft Ozone ISA, for short-term total mortality and CV effects, the causality determinations were downgraded from *likely* to *suggestive*, based on the studies reviewed in the Draft Ozone ISA, which were limited as indicated above. Part of the rationale for downgrading these causality determinations was continued limitations in the epidemiological evidence. An open question is whether that causality determination would have been downgraded had all the evidence been considered. This needs to be addressed in the Draft Ozone ISA, with a broadening of the epidemiology review criteria, and re-assessment of the strength of the causality relationship for these categories of health effects. The EPA has provided a list of the studies of short-term ozone exposure and cardiovascular morbidity and mortality that were excluded from the ISA review because of the geographical location of the study. The CASAC recommends that these studies be included in the EPA's review and that the causality determinations for short-term ozone exposure and mortality, and short-term ozone exposure and cardiovascular effects, be reconsidered with these additional studies as part of the literature set.

Additional studies for inclusion:

- The comments in Appendix B and other comments received from the non-CASAC member consultants and the public identify some omitted relevant studies (e.g., Moore et al., 2012).
- Page 3-14, last paragraph, add Frampton et al. (2015) to the list of new studies of lung function effects in the range of 100-300 ppb. This study included both glutathione-S-transferase M1 (GSTM1) sufficient and null subjects and showed no effects of GSTM1 gene status on lung function responses.
- Vitolo et al. (2018), who investigate the associations between mortality and air pollution in a large study of the United Kingdom using a Bayesian network graphical model and big data technologies while considering topography, climate, and regional effects.

Biological Plausibility

The CASAC encourages the EPA to include relevant positive and negative key studies when mapping biological plausibility pathways, as well as information about exposure concentrations.

In Section 5.1.2 (short-term ozone metabolic effects, biological plausibility) the EPA needs to distinguish better between short-term and long-term effects of ozone on metabolism. It seems that this section is attributing short-term ozone exposures to diabetes development. But how single perturbations will predispose to chronic disease needs more discussion. Similarly, an important question is whether a brief stress response, in the absence of symptoms or other consequences, constitutes an adverse health effect. This could be considered a physiological response to a variety of stimuli. For example, it can occur in response to exercise.

Figure 3-1 provides an excellent synthesis of known and suspected biological pathways mediating ozone respiratory health effects. Some suggestions for further refinement:

- Altered heart rhythm is included here, which is obviously not strictly a respiratory response. But other non-respiratory links are not included here that are consequences of autonomic nervous system modulation and stress responses, including systemic inflammation and metabolic processes. This seems to be an inconsistency. The CASAC suggests removing altered heart rhythm from Figure 3-1 for consistency.
- Impaired host defense is shown linked solely with oxidative stress, but there is evidence for other likely contributing pathways, including airway injury, morphologic airway changes, and stress responses (elevated cortisol). The CASAC suggests moving this box one column to the right, ungroup it from morphologic changes and allergic responses and show as one of the downstream effects.
- The pathway indicating that adrenal effects mediate airway injury/inflammation is based on a single study in rats (Miller et al., 2016b). This finding runs counter to physiologic expectations (adrenal-mediated stress response would be expected to follow acute inflammation/injury, not mediate it) and there is no evidence to support that this occurs in humans. Without further confirmation in additional studies or other species, or support of this directionality in humans, the CASAC suggests making this line dotted.

Additional Comments for Appendices 3 - 7

3.1.4.1.1.2 Cigarette Smoking, P. 3-18. This section summarizes the Bates et al. (2014) study as showing similar lung function responses between smokers and nonsmokers and indicates that this finding differs from previous studies. But the smokers in the Bates study were so-called “light” smokers, on average smoking about half of a pack per day (ppd) for 6 years, for a total of 3 pack-years. This likely explains the difference from prior studies, which involved subjects with greater tobacco use, and this should be noted in the summary. For example, in Frampton et al. (1997), one of the studies demonstrating significantly reduced lung function effects in smokers compared with never-smokers, only smokers of at least 1 ppd for a minimum of 3 years were included. The mean pack-years of smoking was 12.8. It should also be noted that, while ozone-induced lung function decrements are attenuated in smokers, lung inflammation is not (Torres et al., 1997), and oxidative stress may actually be increased (Voter et al., 2001). This is an example of a situation where adverse respiratory effects of ozone may be occurring in the absence of lung function changes.

3.1.4.2.2 Animal Toxicological Studies, p. 3-23. Symptoms by definition are self-reported, and animals are obviously unable to report symptoms. It should be more clearly pointed out here that symptoms cannot be assessed in studies of rodents. Cough, or any other change in respiratory status, when reported by an observer, is a sign or an observation, not a symptom. It is only a symptom when reported by the individual experiencing it.

3.1.4.4.4 Integrated Summary for Respiratory Tract Inflammation, Injury, and Oxidative Stress, P. 3-38, line 1-2. Change "FEV_1" to "FVC." FEV_1 is affected by changes in both volume (FVC, restrictive) and airways obstruction (FEV_1/FVC).

3.1.5.4.2 Animal Toxicological Studies, P. 3-46, line 8. "These effects include sensory and pulmonary irritation…" The distinction here between "sensory" vs "pulmonary" irritation does not make sense. Pulmonary irritant responses have major sensory components. This phrase appears to be taken straight from the Hansen et al. (2016) abstract, but the terminology used in that abstract is not reflective of airway physiology. Sensory vagal-mediated inputs are important throughout the respiratory tract. The upper-lower airway distinction here is incorrect, and is irrelevant to the point being made in this summary. The Hansen et al. (2016) study examines pulmonary outcomes, not upper airway responses.

Table 4-4. The study by Frampton et al. (2015) did not assess left ventricular developed pressure (LVDP). The cardiac function outcomes were cardiac index, stroke-volume index, and left ventricular ejection time. It is perhaps worth mentioning that these measures were obtained via impedance cardiography, rather than directly or via echocardiography.

Table 4-19. The study of Rich et al. (2018) measured systolic blood pressure (SBP) as well as diastolic blood pressure (DBP).

Table 4-26. This table should include Frampton et al. (2017), which examined a number of coagulation parameters, without significant effects.

Table 4-29. Add Frampton et al. (2017) here as well.

5.1.5.2 Liver Outcomes, p. 5-14: "Acute-phase liver proteins, such as [C-reactive protein] CRP, can act as sensors of liver function." This is not accurate. CRP is made in the liver and is a marker of systemic inflammation. Its production is driven by interleukin-6, released by a variety of cells during inflammation. Although CRP is produced in the liver, it is not considered a clinically useful marker of liver function.

5.1.5.3.4 Summary, p. 5-17. "Elevated circulating stress hormones are consistently observed in animal models and in controlled human exposure studies after short-term ozone exposure." This should be "in a single human controlled exposure study." In addition, the last sentence of this summary statement ("Thus, neuroendocrine stress activation is essential to the development of adverse metabolic outcomes after short-term ozone exposure.") is overly broad and not completely supported by the described (adrenalectomy) studies.

5.1.5.4 Serum Lipids, p. 5-18. The description of the Chen et al. (2016a) study is unclear, and it seems incorrect. According to the abstract, this study deals with changes in lung function and nasal inflammation among schoolchildren. Was the reference intended to be Chen et al. (2016b)?

5.1.8. Summary and Causality Determination, p. 5-23. Ketone bodies as a "biomarker" of diabetes is not accurate. Ketone bodies are also a biomarker of starvation or consuming a ketogenic (low carb) diet. It is more accurately a marker of metabolic stress or perturbation with regards to glucose utilization. Ketone bodies increase with diabetic ketoacidosis and can be considered a marker of that acute condition, but not of diabetes in general. Transient elevation of ketone bodies does not mean a person has or will get diabetes.

Table 5-1. "Consistent epidemiologic evidence" is inappropriate given there is only one study supporting it.

Table 6-1. Under "Key Evidence" the statement is made that, "Animal toxicological and controlled human exposure studies do not provide consistent evidence of potential biological pathways." Actually, the experimental animal evidence for CV effects is fairly robust and convincing. It was mostly the inconsistency in the human studies and the relative lack of CV morbidity studies that led to the change in causality determination.

7.2 Nervous system effects. Apparently included in this are the effects on the pulmonary irritant receptor/autonomic pathways that are well-established pulmonary effects in both animals and humans. Consideration should be given to separating this, and having this section include effects beyond the pulmonary irritant response loop, perhaps limiting it to brain, cognitive, and behavioral effects. Otherwise this category would be considered causal based on the known local pulmonary neurological effects.

7.2.2.7 Summary and Causality Determination, p. 7-42, line 21, "reproductive" effects should presumably be "nervous system" effects here.

Appendix 8

Please comment on the identification, evaluation and characterization of the available scientific evidence from studies of ecological effects of ozone, and the application of information from these studies, as presented in Appendix 8, to inform causality determinations for these welfare outcomes.

Determinations are made about causation by evaluating evidence across scientific disciplines and are based on judgements of consistency, coherence and biological plausibility of observed effects, as well as related uncertainties. It is noted that the Draft Ozone ISA uses a formal causal framework to classify the "weight of the evidence" using a five-level hierarchy that characterizes the evidence that forms the basis of causality determinations for welfare effect categories of a "causal relationship" or a "likely to be causal relationship" or describes instances where a causality determination has changed (i.e., "likely to be causal" changed to "suggestive of, but not sufficient to infer a causal relationship"). Other relationships between ozone and welfare effects include "suggestive of, but not sufficient to infer" and "inadequate."

There are 12 causality determinations for ecological effects of ozone that are generally organized from the individual-organism scale to the ecosystem scale presented in Figure ES-5 in the Draft Ozone ISA. To summarize the findings of the 2013 Ozone ISA (U.S. EPA, 2013), five are causal relationships (i.e., visible foliar injury, reduced vegetation growth, reduced crop yield, reduced productivity, and altered below ground biogeochemical cycles), and two are likely to be causal relationships (i.e., reduced carbon sequestration and altered ecosystem water cycling). One of the endpoints, alteration of terrestrial community composition, has now been concluded to be a "causal relationship" wherein the 2013 Ozone ISA this endpoint was classified as "likely to be causal." Three new endpoint categories (i.e., increased tree mortality, alteration of herbivore growth and reproduction, alteration of plant-insect signaling) not evaluated in the 2013 Ozone ISA are all determined to have a "likely to be causal relationship" with ozone. Plant reproduction, previously considered as part of the evidence for growth effects is now a stand-alone causal relationship as illustrated in Figure ES-5.

Visible foliar injury from ozone exposure has been well characterized and documented over decades involving many trees, shrubs, herbaceous and crop species in using both long-term field studies and laboratory approaches. Even more recent experimental evidence continues to show consistent association between visible injury and ozone exposure supporting a "causal relationship" between ozone and visible foliar injury. Consistent with the 2013 Ozone ISA, there is a "causal relationship" between ozone and reduced plant growth and a "causal relationship" between ozone and reduced crop yield and quality. In the 2013 Ozone ISA, the EPA considered reproduction in the same category with plant growth. Increased information of plant reproduction (such as flower number, fruit number, fruit weight, seed number, rate of seed germination) and evidence for direct negative effects on reproductive tissues, as well as for indirect negative effects (resulting from decreased photosynthesis and other whole plant physiological changes) warrants a special causality determination of a "causal relationship" between ozone exposure and reduced plant reproduction. Since the 2013 Ozone ISA, large-scale statistical analysis of many factors concluded that county-level ozone concentrations averaged over the study period significantly increased tree mortality and many plant functional types. This evidence, combined with observations of long-term declines of conifer forests in several high ozone regions and new experimental evidence that sensitive genotypes of, particularly, aspen trees have increased mortality with ozone exposure, support a "likely to be causal relationship" between ozone exposure and tree mortality.

In addition to the direct effects of ozone on plants, ozone can alter ecological interactions between plants and other species, including herbivores that may consume ozone-exposed vegetation. Some recent evidence of insect herbivores in previous ozone assessments and new studies covering a range of species provide collective evidence that supports a "likely to be causal relationship" between ozone exposure and altered herbivore growth and reproduction. Many plant-insect interactions are mediated between volatile plant signaling compounds, which plants use to signal other members within an ecological community. New evidence from multiple studies show altered/degraded emissions of chemical signals from plants and reduced detection of plant signaling compounds by insects. Therefore, the collective evidence supports "a likely to be causal relationship" between ozone exposure and alteration of plant-insect signaling.

At the ecosystem scale, ozone causes suppression of plants' photosynthesis which can lead to reduced ecosystem carbon content. Consistent with the conclusions of the 2013 Ozone ISA, there is a "causal relationship" between ozone exposure and reduced productivity and a "likely to be causal" relationship between ozone and reduced carbon sequestration. Recent evidence continues to support a "causal

relationship" between ozone exposure and the alteration of below ground biogeochemical cycles. Ozone can affect water use in plants through several mechanisms and ultimately affect plant evapotranspiration, which may in turn lead to possible effects on hydrogeological cycling. Evidence continues to support the conclusion of the 2013 Ozone ISA that there is a "likely to be causal relationship" between ozone and alteration of ecosystem water cycling. Alteration of community composition of some ecosystems, including conifer forests, broadleaf forests and grasslands, and altered fungal and bacterial communities in the soil reported in the 2013 Ozone ISA is augmented by additional evidence for effects in forests and grassland communities indicating a change in the causality determination to a "causal relationship" between ozone exposure and altered terrestrial community composition of some ecosystems.

The causality determinations for ecological effects are summarized as follows:

- Conclusions from the 2013 Ozone ISA that support the seven conclusions of causality in the current Draft Ozone ISA include: 1) visible foliar injury; 2) reduced vegetation growth; 3) reduced plant reproduction; 4) reduced yield and quality of agricultural crops; 5) reduced productivity in terrestrial ecosystems; 6) alteration of below ground biogeochemical cycles; and 7) alteration of terrestrial community composition.

The "weight of the evidence" appears to strongly support the previous conclusions from the 2013 Ozone ISA subsequently identified in the conclusions in the current Draft Ozone ISA. The summary of five causality determinations for ecological effects in the current Draft Ozone ISA, which build on the conclusions from the 2013 Ozone ISA, include the following:

1. Reduced plant reproduction from no "separate causality" to a "causal relationship" with ozone exposure;
2. Increased tree mortality "causality not assessed" and changed to "likely to be a causal relationship;"
3. Alternation of herbivore growth and reproduction changed from "causality not assessed" to "likely to be causal relationship;"
4. Alteration of plant-insect signaling "causality not assessed" changed to "likely to be a causal relationship;"
5. Alteration of terrestrial community composition changed from "likely to be a causal relationship" to "causal relationship."

For these five causality determinations for ecological effects that have changed in terms of conclusions in the current Draft Ozone ISA from the conclusions from the 2013 Ozone ISA will be more fully evaluated in terms of preliminary comments from the initial review of these data.

Appendix 8 "Ecological Effects" in the current Draft Ozone ISA evaluates the relevant scientific information on ecological effects as part of the review of the air quality criteria for ozone and other photochemical oxidants and to help form the scientific foundation for the review of the secondary NAAQS for ozone. This Appendix serves as an update to Chapter 9 of the 2013 Ozone ISA. The majority of the evidence for ecological effects is for vegetation. Effects at the individual plant level can result in broad ecosystem-level changes, such as productivity, carbon storage, water cycling, nutrient cycling, and community composition. The current Draft Ozone ISA adopts the use of the PECOS tool to further define the scope of the current review by conveying the criteria for inclusion or exclusion of studies. The units of study as defined in the PECOS tool for ecological effects of ozone are the

individual organism, species, population, community, or ecosystem. It should be noted that all studies included in the current Draft Ozone ISA were conducted at concentrations occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations observed in the United States. For ecological endpoints for which the 2013 Ozone ISA concluded that the evidence was sufficient to infer a causal relationship (i.e., foliar injury, vegetation growth, ecosystem productivity, yield and quality of agricultural crops, below ground biogeochemical cycling), these are fully evaluated in the current Draft Ozone ISA. In terms of new causal determinations or a change in causal determination from the 2013 Ozone ISA, the following causality determinations for ecological effects of ozone will be addressed in the current review. At the community level, biodiversity in terms of terrestrial community composition is now "causal," and species interactions including plant-insect signaling is a new determination and "likely to be causal." In addition, tree survival is changed to "likely causal" and growth of insect herbivores feeding on ozone-affected plants is "likely causal." The plant reproduction endpoint is now separate from plant growth and a new determination as "causal" and new determination of growth and reproduction is "likely to be causal" is assigned to insect herbivores affected by ozone. All causality determinations or changes in causality determination from the 2013 Ozone ISA will be thoroughly considered in the present series of comments. The current review only evaluates studies conducted in North America. In the PECOS tool for ecological effects, relevant study designs include laboratory, greenhouse, field, gradient, open top chamber (OTC), free air carbon dioxide enrichment (FACE), and modeling studies.

Visible Foliar Injury in Biomonitoring

In the 2013 Ozone ISA, the evidence was sufficient to conclude that there is a causal relationship between ambient ozone exposure and the occurrence of ozone-induced visible foliar injury on sensitive plant species across the United States. Visible foliar injury from exposure to ozone has been well characterized and documented on many tree, shrub, herbaceous, and crop species through research beginning in 1958. Ozone-induced visible foliar injury is considered diagnostic because it has been experimentally induced and it is considered a bioindicator for ozone exposure in plants. As described in the PECOS tool, the scope for new evidence reviewed in the section limits studies to those conducted in North America at concentrations occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations. Experimental evidence continues to show a consistent association between visible injury and ozone exposure in plants. Since the 2013 Ozone ISA, several studies have further characterized modifying factors: 1) additional field studies have shown dry periods tend to decrease the incidence and severity of ozone-induced visible foliar injury; 2) data used in additional species from greenhouse studies add to the evidence that sensitivity to ozone varies by the time of day in plants; 3) phenotypic variation of foliar sensitivity to ozone has been observed; 4) in OTC exposure (mean 12-hour ozone concentration of 37 ppb for 118 days) foliar injury to loblolly pine seedlings were not related to seedling inoculation with root-infecting fungi (Chieppa et al., 2015).

Since the 2013 Ozone ISA, several additional studies have been conducted on bioindicator species:

1. Cutleaf coneflower is an ozone bioindicator species native to Great Smokey Mountains National Park;
2. Tree of heaven, an established invasive species found widely across the United States, has been identified as an effective ozone bioindicator species by the National Park Service and Forest Service (Smith et al., 2008; Kohut, 2007).

In greenhouse exposures, foliar injury occurred at 8-hour average ozone exposure levels of 60 – 120 ppb with greater injury corresponding to higher exposure (Seiler et al., 2014). As noted in the 2013 Ozone ISA, visible foliar injury usually occurs when sensitive plants are exposed to elevated ozone concentrations in a predisposing environment. A major modifying factor for ozone-induced visible foliar injury is the amount of soil moisture available to a plant during the year that the visible foliar injury is being assessed. This is because the lack of soil moisture generally decreases stomatal conductance of plants and, therefore, limits the amount of ozone entering the leaf that can cause injury. Visible foliar injury from ozone exposure has been well characterized for decades using both long-term field studies and laboratory approaches. Since the 2013 Ozone ISA, new research on bioindicator species and the further characterization of modifying factors have provided further support for the effects. New information is consistent with the conclusions of the 2013 Ozone ISA that the body of evidence is sufficient to infer a causal relationship between ozone exposure and visible foliar injury. With the decades of research, both in field observation as well as experimental studies related to the foliar injury endpoint, the body of evidence remains very strong to infer a causal relationship between ozone exposure and visible foliar injury.

Plant Growth

In the 2013 Ozone ISA, the evidence was sufficient to conclude that there is a causal relationship between ambient ozone exposure and reduced growth of native woody and herbaceous vegetation. In the 2013 Ozone ISA, it was concluded there is strong and consistent evidence that exposure to ozone decreases photosynthesis and growth in numerous plant species. The evidence available at that time and now discussed in the current Draft Ozone ISA shows that ambient ozone concentrations causes decreased growth (measured as biomass accumulation) in annual, perennial, and woody plants inclusive of crops, annuals, grasses, shrubs, and trees. A meta-analysis by Wittig et al. (2009) found that the average ozone exposures of 40 ppb significantly decreased annual total biomass by 7% across 263 studies. Biomass declines were linked to reductions in photosynthesis (U.S. EPA, 2013), which are consistent with cumulative plant uptake of ozone into the leaf (Wittig et al., 2007). Further, there is evidence ozone may change plant growth patterns by significantly reducing carbon allocated to roots in some species. Since the 2013 Ozone ISA, there is more evidence from experimental studies that support detrimental effects of ozone on plant growth:

1. Results from aspen-only stand at the Aspen FACE experiment in Wisconsin showed a decrease of 12 – 19% in the relative growth rate of 3 of 5 genotypes of aspen studied;
2. When site-level results from Aspen FACE experiment were scaled up using the forest landscape model (LANDIS II), ozone was found to significantly reduce landscape biomass;
3. A meta-analysis of 9 studies examining intra-specific variation in juvenile tree growth under elevated ozone, found that elevated ozone generally reduced photosynthetic rate as well as height growth and stem volume;
4. A study using the invasive Chinese tallow tree suggests ozone response may be genotype-specific;
5. Model simulations coupled with established U.S. EPA ozone exposure response functions in seedlings, estimated relative biomass loss at 2.5% for Ponderosa pine and 2.9% for aspen; and
6. In another estimation of biomass loss of adult trees across the United States for modeled ozone values, eastern cottonwood and black cherry showed high sensitivity.

In addition to these studies, there is a recent global scale synthesis of published ozone exposures studies that document reductions in biomass due to ozone exposure in over 100 plant species (Bergmann et al., 2017). In the current Draft Ozone ISA, there is strong scientific evidence sufficient to conclude that there is a causal relationship between ambient ozone exposure and reduced growth of native woody and herbaceous vegetation.

Reduced Plant Reproduction

In the 2013 Ozone ISA, reduced plant reproduction was not separated from plant growth for causality determination. However, in the current Draft Ozone ISA, reduced plant reproduction is scientifically defended for a causal relationship between plant reproduction metrics and exposure to ozone. In fact, the recent literature shows that across most plant reproduction metrics (such as flower number, fruit number, fruit weight, seed number, and rate of seed germination) with elevated exposure concentrations that ozone has significant negative effects on plant reproduction. In a first of its kind study, Leisner and Ainsworth (2012) conducted a quantitative meta-analysis to assess the general magnitude and direction of the effects of ozone exposure on plant reproduction. In experiments that used ambient air as the control, average fruit weight decreased 51% (at an average exposure of 98 ppb), which was the largest effect observed in this part of the meta-analysis, and seed number decreased approximately 10% (at an average exposure of 68 ppb). In studies with ozone-sensitive species of clover, Sanz et al. (2016) showed that reproduction was reduced significantly with increasing ozone exposure. Gillespie et al. (2015) isolated the effects of ozone on particular reproductive tissues of tomato. Pollen grains exposed to ozone have significantly reduced germination and pollen tube growth in vitro. Reductions in pollen viability is an extremely important plant reproduction metric.

Timing of ozone exposure relative to reproductive development stages can affect reproductive outcomes in some cases. Flowers exposed to ozone early in their development tended to produce shorter fruits than flowers exposed later in their development. There appears to be adequate information, particularly from the quantitative meta-analysis reported by Leisner and Ainsworth (2012) supporting a causal relationship between ozone exposure and reduced plant reproduction. The strength of the scientific support for supporting a "causal relationship" is not as strong as with visible foliar injury and reduced vegetation growth. However, with the separate category of reduced plant reproduction, it can be concurred that causality does exist between ambient ozone exposure and this plant metric. It has been shown that diverse metrics of plant reproduction decline under ozone concentrations occurring in either the environment or under experimental conditions within an order of magnitude of recent concentrations. Metrics of plant reproduction, fruit number and fruit weight, show reductions under increased ozone when combined across species for ozone concentrations that span 40 to >100 ppb. Finally, experimental ozone exposure at multiple experimental settings (such as in vitro, whole plants in the laboratory, whole plants and/or reproductive structures in the greenhouse, and whole plant communities in field settings) convincingly show ozone independently reduces plant reproduction. The CASAC concurs with the EPA conclusion that previous evidence and the new evidence is sufficient to infer a "causal relationship" between ozone exposure and reduced plant reproduction.

Plant Mortality

In the 2013 Ozone ISA, causality was not assessed for increased tree mortality involving ozone exposure. The conclusion in the current Draft Ozone ISA is that there is "likely to be a causal relationship" between ozone exposure and plant mortality. Several new studies examine the impacts of

ozone exposure on plant mortality that included the fraction of individuals in a population that die over a given timeframe. These experiments were focused on tree species demonstrating ozone exposure can affect tree mortality. For instance, in the Aspen FACE experiment, the survival of sensitive aspen genotypes 271 and 259 declined significantly between 1997 and 2008 under elevated ozone exposures (Moran and Kubiske, 2013). In addition, Dietz and Moorcroft (2011) conducted a large-scale analysis of factors contributing to annual mortality of trees and functional types in the forests of the eastern and central United States. In their analysis, ozone was ranked 9th on a list of 13 factors that forests were sensitive to ozone's effects with a similar magnitude to that of precipitation. Mortality in 8 out of 10 plant functional types were significantly correlated with ozone 8-hour max exposures. Therefore, studies of tree mortality indicate that ozone affects this endpoint. Studies linking ozone and tree mortality are consistent with known and well-established individual plant level mechanisms that explain ozone phytotoxicity, including variation and sensitivity and tolerance based on age class, genotype, and species. Experimentally elevated ozone exposures have been shown to increase mortality in sensitive Aspen genotypes. Considering the previous evidence and new evidence reviewed in the current Draft Ozone ISA, it is sufficient to infer a "likely to be causal relationship" between ozone exposure and tree mortality.

Reduced Crop Yield and Quality

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a "causal relationship" between ozone exposure and reduced yield and quality of agricultural crops. The detrimental effect of ozone on crop production has been recognized since the 1960s and there is a large body of research that has subsequently characterized decreases in yield and quality of agricultural crops. As described in the PECOS tool, the scope of new evidence reviewed in this section are limited to studies conducted in North America at ozone concentrations occurring in the environment or experimental ozone concentrations within an order of magnitude of research concentrations.

For soybeans, additional studies in Illinois report decreased seed/crop yield (Leisner et al., 2017). A linear decrease in soybean yield was observed across two growing seasons at the rate of 37-39 kg/ha per ppb cumulative ozone exposure over 40 ppb. For wheat, meta-analysis using data from the United States and other countries provide further supporting evidence that current levels of ambient ozone decrease growth, quality, and yield (Pleijel et al., 2018). New studies in non-soybean legumes include evaluation of biomass and seed yield in ozone-exposed snap bean under high- and low-vapor pressure deficit conditions (Fiscus et al., 2012). U.S. modeling studies in the 2013 Ozone ISA found that ozone generally reduced crop yield and that different crops showed different sensitivity to ozone (Avnery et al., 2011). Newly available regional and national scale analyses of ozone effects on major crops in the United States, including soybean, wheat, and maize have further enabled characterization and quantification of yield losses (McGrath et al., 2015).

The relationship between ozone exposure and reduced crop yield is well established in the scientific literature and continues to be an active area of research with many new scientific papers being published since the 2013 Ozone ISA. Recent advances in characterizing ozone effects on U.S. crop yield include further geographic and temporal refinement of ozone sensitivity in national scale estimates of maize and soybean losses from ozone based on actual yield data. The new scientific information published is consistent with the conclusions of the 2013 Ozone ISA that the body of evidence is sufficient to infer a "causal relationship" between ozone exposure and reduced yield and quality of agricultural crops.

Herbivores: Growth, Reproduction, and Survival

In the 2013 Ozone ISA, there was no causality determination between ozone exposure and effects on herbivores. Ozone exposure can lead to changes in plant physiology, such as by modifying the chemistry and nutrient content of leaves. These changes can have significant effects on herbivore physiology and behavior. There was no consensus in the 2013 Ozone ISA on how insects and other wildlife respond to elevated ozone. Since that review, additional research has been published for more herbivorous insects as well as a few mammalian herbivores at various levels of ozone exposure. As described in the PECOS tool, the scope of this review includes studies in which alterations in invertebrates and vertebrate responses were measured in individual species or at the population and community levels as related to concentrations of ozone occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations. The 2013 Ozone ISA included a meta-analysis that included 16 studies published on insect herbivore species between 1996 and 2005 found that elevated ozone decreased development time and increased pupil mass in insect herbivores with more pronounced effects occurring with longer durations of ozone exposure (Valkama et al., 2007). Since the 2013 Ozone ISA, there is new evidence for endpoints related to growth, reproduction, and survival in insect herbivores encompassing the orders Coleoptera, Hemiptera, and Lepidoptera. With the available science reported in the current Draft Ozone ISA regarding the effects of ozone on growth, reproduction, and survival of, particularly, insect herbivores, substantial new information has been made available in order to assess a causality relationship. In addition, population and community level responses reveal that changes in host–plant quality resulting from elevated ozone can alter the population density and structure of associated insect herbivore communities ultimately affecting ecosystem processes (Cornelissen, 2011). Recent studies reviewed in the current Draft Ozone ISA include multiple experimental studies conducted by many research groups that expand the evidence base for the effects of elevated ozone on growth and reproduction in herbivores. It is recognized that while effects were observed there remains a more limited number of studies on the effects of ozone on survival and population/community level responses. Recognizing that since the 2013 Ozone ISA and with increased research efforts on herbivore response to plants impacted by ozone, a new causality determination appears justified that the body of evidence is sufficient to infer a "likely to be causal" relationship between ozone exposure and alteration of herbivore growth and reproduction.

Alteration of Plant-Insect Signaling

In the 2013 Ozone ISA, there was "no causality" determination between ozone exposure and alteration of plant-insect signaling. Plants signal to other ecological community members through the emission of volatile plant signaling compounds (Blande et al., 2014). Each signal emitted by plants has an atmospheric lifetime and a unique signature comprised of different ratios of individual hydrocarbons that are susceptible to atmospheric oxidants, like ozone (Yuan et al., 2009). Insects and other fauna discriminate between chemical signals of different plants. As described in the PECOS tool, the scope in the current Draft Ozone ISA for considering plant-insect signaling include studies that assess altered plant-insect signaling in response to concentrations of ozone occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations. Under conditions of elevated ozone, the degradation of plant signaling compounds resulted in bumble bees orienting significantly less towards floral scent queues and exhibiting preference for artificial flowers closer to the ozone source (Farré-Armengol et al., 2015). As reported previously, herbivorous insects use plant signaling compounds to locate suitable host plants and ozone can alter these interactions (Blande et al., 2010). In chamber studies, elevated ozone reduced the ability of insect herbivores to find their plant

host (Li et al., 2016). Striped cucumber beetles could not distinguish between clean air and air containing floral volatiles when the ozone concentration exceeded 80 ppb (Fuentes et al., 2013). In addition, plant defense responses include emission of plant-signaling compounds to attract predators and parasitoids that target the herbivores feeding on the plant. In studies reviewed in the 2013 Ozone ISA and new studies on parasitoid-host attraction show either reduced, enhanced, or unaffected behavior by elevated ozone (Cui et al., 2016). Altered plants signaling to natural enemies of herbivores disrupts predator-prey trophic interactions. The interaction of ozone (>50 ppb) with plant signaling compounds disrupts the production, emission, dispersion, and lifespan of these compounds. Considering the available evidence reported in the 2013 Ozone ISA and more recent research efforts while as well recognizing uncertainties around how chemical signaling responses observed in the laboratory translate to natural environments, the current Draft Ozone ISA makes a new causality determination that the body of evidence is sufficient to infer a "likely causal relationship" between ozone exposure and alteration of plant-insect signaling.

Reduced Productivity in Terrestrial Ecosystems

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a causal relationship between ozone exposure and reduced plant productivity. The terrestrial carbon cycle integrates processes at various scales, ranging to organelles to individuals to biomes (Chapin et al., 2002). Gross primary productivity, which is the influx of CO_2 from the atmosphere via photosynthesis at the ecosystem scale is fundamental to global carbon cycling. Since the 2013 Ozone ISA, two new studies have reported on the effects of ozone on gross primary productivity. Fares et al. (2013) conducted statistical analysis of data to quantify the effect of ozone on carbon assimilation. In California, ozone decreased carbon assimilation by 12% in pine forests in the Sierra Nevada and by 19% in an orange grove in the Central Valley. Yue and Unger (2014) adopted the same ozone-damaged thresholds in their analysis that were used in previous models to assess ozone damage. Decreases in gross primary productivity as a result of ozone range from 1-14% and were greatest at sites showing both high stomatal conductance and high growing season ozone concentrations. Carbon assimilated into plant tissue via photosynthesis is either respired or contributes to net primary productivity, which is often measured as the rate of plant biomass accumulation. While much of the research published since 2013 Ozone ISA is confirmatory, other work has provided new mechanistic insight into the effects of ozone on net primary productivity. Evidence of the effect of ozone exposure in ecosystem productivity comes from many different experiments with different study designs in a variety of ecosystems and models. New information is consistent with conclusions of the 2013 Ozone ISA that the body of evidence is sufficient and increasing to infer a "causal relationship" between ozone exposure and reduced ecosystem productivity.

Reduced Carbon Sequestration in Terrestrial Ecosystems

Terrestrial carbon sequestration is the sum of carbon contained within biomass and soil within a defined ecosystem typically quantified on a multi-year scale (Körner, 2006). As in the 2013 Ozone ISA, most assessments of the effects of ozone on terrestrial carbon sequestration are from model simulations. However, an assessment was done of the effect of ozone on ecosystem carbon content at the Aspen FACE experiment (Talhelm et al., 2014). At the conclusion of the Aspen FACE experiment after 11 years of fumigation, it was observed that elevated ozone decreased ecosystem carbon content (in plant biomass, litter, and soil carbon to 1 m in depth) by 9%. Total tree biomass carbon was 15% lower under elevated ozone with decreased woody biomass counting for nearly all of the effect of tree biomass. The results from the Aspen FACE experiment and the model simulations provide further evidence that ozone

can decrease ecosystem carbon sequestration. Although the decrease in net primary productivity were temporary in the Aspen FACE experiment, the 10% decrease in cumulative net primary productivity at Aspen FACE was associated with a 9% decrease in ecosystem carbon storage (Talhelm et al., 2014). The relationship between ozone exposure and terrestrial carbon sequestration is difficult to measure at the landscape scale. Most of the evidence regarding this relationship is from model simulations, although this endpoint was examined in a long-term manipulative chamber-less ecosystem experiment known as Aspen FACE, already described. Even with limitations, the result from the Aspen FACE experiment and supported by model simulation provide further evidence that is consistent with the conclusions of the 2013 Ozone ISA that the body of evidence is sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and reduced carbon sequestration in ecosystems.

Soil Biogeochemistry

The 2013 Ozone ISA concluded there is a "causal relationship" between ozone exposure and the alteration of below ground biogeochemical cycles. This causality determination was based on the body of evidence known at that time. The 2013 Ozone ISA presented evidence that ozone alters multiple below ground endpoints, including root growth, soil food web structure, soil decomposer activities, soil respiration, soil carbon turnover, soil water cycling, and soil nutrient cycling. The new evidence since the 2013 Ozone ISA included in the current Draft Ozone ISA confirms ozone affects soil decomposition, soil carbon, and soil nitrogen. Soil carbon is often a mix of inorganic and organic forms of carbon, the latter may be from living and/or dead plant animal, fungal, and bacterial organisms. The effects of ozone on several aspects of soil carbon have been investigated. Ozone can alter the cycling of nitrogen in the soil via its direct effect on plants. Nitrogen is an important element to plant life as it is often the limiting nutrient from most temperate ecosystems. The 2013 Ozone ISA documented mixed results of ozone effects on soil nitrogen pools and processes with results indicating no effect in meadow nitrogen biomass or potential nitrification and denitrification (Kanerva et al., 2006). While ozone was shown to increase nitrogen released from litter in a forest (Stoelken et al., 2010), ozone decreased gross nitrogen mineralization (Holmes et al., 2006) at Aspen FACE and nitrogen release from soil litter. The 2013 Ozone ISA presented evidence that ozone was found to alter multiple below ground endpoints, including root growth, soil food web structure, soil decomposer activities, soil respiration, soil carbon turnover, soil water cycling, and soil nutrient cycling. New evidence since the 2013 Ozone ISA included in this assessment confirms ozone effects on soil decomposition, soil carbon, and soil nitrogen. Overall, the evidence does not change the conclusions from the 2013 Ozone ISA and, therefore, suggests that ozone can alter soil biogeochemical cycling of carbon and nitrogen, although the direction and magnitude of these changes often depends on the species, site, and time of exposure. Currently, it is recognized that it does not appear to be a consistent exposure-response relationship. The body of evidence is sufficient to conclude that there is a "causal relationship" between ozone exposure and the alteration of below ground biogeochemical cycles.

Alteration of Terrestrial Community Composition

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and alteration of terrestrial community composition of some ecosystems. Ozone altered above ground plant communities, such as conifer forests, broadleaf forests, and grasslands and altered fungal and bacterial communities in the soil in both natural and agricultural systems. Ozone effects on individual plants can alter the larger plant community as well as the below ground community of microbes and invertebrates, which depend on plants as carbon sources. In the

2013 Ozone ISA, evidence of ozone effects on forest composition was drawn from the observational studies of conifer decline correlated with ozone exposure (Allen et al., 2007). New evidence suggests that ozone alters tree competitive interactions for nutrients, such as consistent with previous research on altered tree community composition at Aspen FACE showed that elevated ozone altered the relative competition for nutrients among aspen genotypes (Zak et al., 2012). Since the 2013 Ozone ISA, new studies extend the scope of evidence regarding forest community composition to include synthesis and models. In the 2013 Ozone ISA, there was evidence of ozone effects on grassland community composition in controlled experimental exposure studies, in models, and in reviews. Key new studies include experimental ozone exposures that allow evaluation of ozone effects on grassland community composition and analyses that explicitly include environmental or annual heterogeneity.

Even with microbes, the 2013 Ozone ISA documented effects of ozone on soil microbial communities with changes in proportions of bacteria or fungi as a result of experimental ozone exposures in grassland mesocosms, peatland mesocosms, and forest mesocosms. In addition, changes in soil microbial communities in agricultural systems was reported (Chen et al., 2010). Even with bacteria, the 2013 Ozone ISA found decreases in bacterial abundance in response to elevated ozone in meadows and forests mesocosms. There have been many new studies reported to assess the effect of elevated ozone on soil bacteria. The 2013 Ozone ISA found effects of ozone exposure on soil fungi. Studies found that ozone exposure decreased fungal biomass in meadow mesocosms, marginally increased fungal abundance in peatland mesocosms and altered fungal community composition in forest soils. Many new studies have evaluated the effects of ozone on fungi since the 2013 Ozone ISA. The 2013 Ozone ISA found evidence sufficient to conclude that there is a "likely to be causal relationship" between ozone exposure and the alteration of community composition of some ecosystems. Evidence of this relationship was presented for forest communities of trees, grassland communities of grasses, herbs, and legumes and soil microbial communities of bacteria and fungi. Recently published papers extend the evidence of each of these topics in the 2013 Ozone ISA.

In forests, previous evidence included correlation on studies across ambient gradients of ozone exposure that found effects of ozone on conifer trees, studies with controlled experimental exposure of trees that found effects of ozone on deciduous trees. Key new studies show that observational and experimental observations of ozone effects on tree species extend to affect regional forest composition in the Eastern U.S. (Wang et al., 2016). In grasslands, previous evidence included multiple studies from multiple research groups to show that elevated ozone shifts the balance among grasses, forests, and legumes. There are new studies that show ozone affected the ratio of grass to legume biomass (Gilliland et al., 2015). In soil microbial communities, previous evidence includes studies that found effects on the ratio of bacteria to fungi in soil communities as well as effects on community composition of mycorrhizal fungi. New studies confirm that elevated ozone alters soil microbial taxa, although as with previous evidence, the strength and directional effects are not consistent across all ecosystems. The 2013 Ozone ISA presented multiple lines of evidence that elevated ozone alters terrestrial community composition, and recent evidence strengthens the understanding of the effects of ozone on plant communities while confirming that the effects of ozone on soil microbial communities are diverse. The body of evidence is sufficient to conclude that there is a "causal relationship" between ozone exposure and the alteration of community composition of some ecosystems.

Alteration of Ecosystem Water Cycling

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and the alteration of ecosystem water cycling. Plants are responsible for part of the ecosystem water cycling through root uptake of soil moisture and groundwater as well as transpiration through leaf stomata to the atmosphere. Changes to this part of the water cycle may in turn affect the amount of water moving through the soil, running off over land or through groundwater and flowing through streams. Ozone can affect water use in plants and ecosystems through several mechanisms, including damage to stomatal functioning and loss of leaf area, which may affect plant and stand transpiration. During the review of the 2013 Ozone ISA, there was debate on the assumption that ozone exposure consistently reduced stomatal conductance in plants. Several studies have found increased conductance, suggesting stomatal dysfunction in response to ozone exposure. However, other studies found ozone caused a loss of stomatal control, incomplete stomatal closure at night, and a decoupling of photosynthesis in stomatal conductance. There is mounting biologically relevant and statistically significant data from multiple studies showing the mechanisms of ozone effects on plant-water use in ecosystem water cycling (reduced leaf area, reduced leaf longevity, changes in root and branch biomass and architecture, changes in vessel anatomy, stomatal dysfunction, reduced sap flow). The most compelling evidence showing effects at the ecosystem level is from studies in Eastern U.S. forests and from the Aspen FACE. All of this new information supports the 2013 Ozone ISA and supports the conclusion in the current Draft Ozone ISA that the body of evidence is sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and the alteration of ecosystem water cycling.

General Comments

1. The CASAC compliments the EPA for the thoroughness and completeness of Appendix 8 as part of the Draft Ozone ISA.
2. The CASAC agrees with the "causal" determinations for the components for ecological effects for: 1) visible foliar injury; 2) reduced vegetation growth; 3) reduced plant reproduction; 4) reduced yield and quality of agricultural crops; 5) reduced productivity in terrestrial ecosystems; 6) alteration of below ground biogeochemical cycles; and 7) alteration of terrestrial community composition. The CASAC agrees with the "likely to be causal" determinations for: 1) increased tree mortality; 2) alteration of herbivore growth and reduction; 3) alteration of plant-insect signaling; 4) reduced carbon sequestration in terrestrial ecosystems; and 5) alteration of ecosystem water cycling.
3. It is the CASAC's impression that a thorough review and reporting of the scientific literature that has been generated since the 2013 Ozone ISA has been incorporated into the current Draft Ozone ISA.
4. In terms of the summary of causality determinations for ecological effects, the CASAC supports the determinations made by the EPA as a function of the available science and its interpretation.
5. Although, historically, the predominant ecological effects assessed with ozone exposure has been with vegetation, Appendix 8, has at least some mention of terrestrial vertebrates, including rabbits, and how they may respond to altered vegetation as a function of ozone exposure. The CASAC thinks that this area should be expanded because alteration of individual plants and plant communities can disrupt terrestrial vertebrates, and not just invertebrates. Therefore, the CASAC recommends consideration of an expanded research plan to look at the implications of altered vegetation communities from ozone exposure and response to terrestrial vertebrate herbivores.

6. Although there is in-depth consideration in other sections of the Draft Ozone ISA involving human health implications from ozone exposure, which are real and well-defined cause and effect relationships that have been scientifically studied a considerable length of time, nothing is mentioned with wildlife. In Appendix 8, there is no mention whatsoever of wildlife toxicology implications for ozone exposure, although human health implications have been considerably considered in other parts of the Draft Ozone ISA. The CASAC recommends to at least consider and develop a research plan for a bird model that could be assessed in terms of the wildlife toxicology of ozone exposure in warm-blooded vertebrates. This would be essentially a "canary in the coal mine" concept for detecting toxic gasses by miners through a bird model. The CASAC thinks this same concept could be implemented utilizing an avian model for the study of ozone exposure in terrestrial warm-blooded non-human vertebrates (Kendall et al., 2010).

Appendix 9

Please comment on the identification, evaluation and characterization of the available scientific evidence from studies of ozone effects on climate, and the application of information from these studies, as presented in Appendix 9, to inform causality determinations for these welfare outcomes.

For effects on climate, changes in the abundance of tropospheric ozone disturbs the radiative balance of the atmosphere by interacting with incoming solar radiation and outgoing longwave radiation. This effect is quantified by radiative forcing, which is the perturbation in net radiation flux at the tropopause caused by a change in radiatively active forcing agent after stratospheric temperatures have readjusted to radiative equilibrium. Through this effect on the earth's radiation balance, tropospheric ozone plays a significant role in the climate system and increases in tropospheric ozone abundance contribute to climate change as addressed in the 2013 Ozone ISA. Recent evidence continues to support a causal relationship between tropospheric ozone and radiative forcing and a "likely to be causal relationship" via radiative forcing between tropospheric ozone and temperature, precipitation and related climate variables referred to as climate change in the 2013 Ozone ISA. New evidence comes from the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) (Myhre et al., 2013) and supporting references. As thoroughly discussed in the current Draft Ozone ISA, none of the new studies indicate a change to either causality determination included in the 2013 Ozone ISA. In terms of effects of tropospheric ozone and climate change, radiative forcing remains a "causal" relationship and temperature, precipitation, and related variables maintain a "likely to be causal" relationship. Consistent with previous estimates in the 2013 Ozone ISA, the current Draft Ozone ISA is consistent with previous estimates, the effect of tropospheric ozone on global surface temperature through its impact on radiative forcing continues to be estimated at roughly 0.1 to 0.3°C since industrial times. While the warming effect of tropospheric ozone in the climate system is established, precisely quantifying changes in surface temperature due to tropospheric ozone changes along with related climate effects requires complex climate simulations. There are current limitations in climate modeling tools that need to be recognized and the need for more comprehensive observational data on these effects represent sources of uncertainty in quantifying the precise magnitude of climate response to ozone changes (Myhre et al., 2013). All of this evidence reinforces the "likely to be causal" relationship between tropospheric ozone and temperature, precipitation, and related climate variables which was referred to as "climate change" in the 2013 Ozone ISA.

General Comments

1. The CASAC compliments the EPA for continuing to clearly characterize and communicate the effects of ozone as related to climate change, building on the 2013 Ozone ISA to the current Draft Ozone ISA.
2. Although evidence has increased supporting the relationship between tropospheric ozone and aspects of climate change, including a "causal relationship" with radiative forcing as well as a "likely to be causal relationship" with impacts on temperature, precipitation, and related climate variables, the causality determinations reached in the 2013 Ozone ISA are even further supported in the current Draft Ozone ISA, and the CASAC strongly concurs with that position.
3. Further research would be useful, particularly quantifying the relationship between regional ozone radiative forcing (RF) and other short-lived climate forcers on the hydrologic cycle, precipitation, and atmospheric circulation patterns; improving understanding of and ability to model critical ozone-climate processes; and continuing exploration of links between precursor pollutant control strategies, climate, and ozone concentrations. These research strategies would be extremely useful as the role of ozone in the climate system scientific arena continues to be better understood and the CASAC recommends that the EPA continue to invest research resources to better define the role of ozone in climate science.

Appendix 10

Appendix 10 provides details on the process by which the draft ISA was developed. Please comment on the usefulness and effectiveness of this appendix. Please provide recommendations on approaches that may improve the communication of the process used to develop the ISA.

The process explained in Appendix 10 appears to be well considered, and parts of it (such as the use of the PECOS tool) appear to be valuable and an advance on earlier approaches. The exposition is, for the most part, clear, although it is written at such a high level that it is hard to determine whether or how the ISA development processes described in Appendix 10 were implemented in practice.

However, as explained in comments on the Integrated Synthesis, the process by which the Draft Ozone ISA was developed has produced results that are not clear to many readers and that appear to have important limitations. The following limitations should be corrected in the final ISA; more importantly for Appendix 10, the process that led to them (or perhaps the implementation of the process, or both) should be improved to avoid such unnecessary limitations in future.

- *Section 10.2 (Literature Search and Initial Screen): Specific criteria for selecting and weighting studies for individual studies and for specific health endpoints are not sufficiently clear* so that even expert readers can understand and reproduce how they were applied. It is also unclear how conclusions would change if consistent criteria were systematically applied for selecting, evaluating, summarizing, and synthesizing studies. International studies and methodology-oriented studies that provide useful recent information on public health effects of changes in ambient ozone levels should probably be included (e.g., Vitolo et al., 2018).
- *Section 10.3 (Study Selection: Full-Text Evaluation of Studies Level 2): Specific criteria for selecting, summarizing, and evaluating studies are not sufficiently clear* so that they can be understood and the results of applying them can be independently reproduced. Spot checks and

public comments suggest that many relevant studies have been excluded (e.g., Moore et al., 2012), and it is not always clear why. Conversely, it is not clear why other studies are interpreted as high quality and capable of providing key evidence (e.g., Tétreault et al., 2016), despite limitations, sometimes clearly stated by the authors, that might seem to preclude clear interpretation of study results. Please see the "Study Inclusion" and "Biological Plausibility" sections in the consensus response to the Appendices 3-7 charge questions as well as individual comments for additional detail on biological information.

In light of these limitations, the CASAC strongly recommends that the EPA work with external experts in causal analysis, biological causation, management science, decision analysis, and risk analysis to revise and improve the current causal determination process. This work should include identifying and adopting techniques for improving group decisions and risk communication under uncertainty, and for reducing biases (e.g., groupthink, confirmation bias, conformation bias, narrow framing, etc.) that frequently undermine the validity of consensus judgments about risk and resulting risk management decisions and policy recommendations. Experts from outside the air pollution health effects area should be included. Much recent and current research from the air pollution health effects community lags by decades other areas of applied science, engineering, epidemiology, and risk analysis in understanding and appropriately applying modern methods and processes of causal analysis, quantitative risk modeling, and management science useful for regulatory risk assessment and science-based risk regulation. The CASAC recommends that the EPA work with the National Academies to identify and use such external expertise to improve the ISA conceptual framework and development process.

References

Allen, E. B., Temple, P. J., Bytnerowicz, A., Arbaugh, M. J., Sirulnik, A. G., Rao, L. E. (2007). Patterns of Understory Diversity in Mixed Coniferous Forests of Southern California Impacted by Air Pollution. *The Scientific World JOURNAL, 7*, 247–263. doi: 10.1100/tsw.2007.72

Arjomandi, M., Balmes, J. R., Frampton, M. W., Bromberg, P., Rich, D. Q., Stark, P., Alexis, N. E., Costantini, M., Hollenbeck-Pringle, D., Dagincourt, N., Hazucha, M. J. (2018). Respiratory Responses to Ozone Exposure. MOSES (The Multicenter Ozone Study in Older Subjects). *American Journal of Respiratory and Critical Care Medicine, 197*(10), 1319–1327. doi: 10.1164/rccm.201708-1613oc

Arjomandi, M., Wong, H., Donde, A., Frelinger, J., Dalton, S., Ching, W., Power, K., Balmes, J. R. (2015). Exposure to medium and high ambient levels of ozone causes adverse systemic inflammatory and cardiac autonomic effects. *American Journal of Physiology-Heart and Circulatory Physiology, 308*(12). doi: 10.1152/ajpheart.00849.2014

Avnery, S., Mauzerall, D. L., Liu, J., Horowitz, L. W. (2011). Global crop yield reductions due to surface ozone exposure: 1. Year 2000 crop production losses and economic damage. *Atmospheric Environment, 45*(13), 2284–2296. doi: 10.1016/j.atmosenv.2010.11.045

Bae, S., Lim, Y.-H., Kashima, S., Yorifuji, T., Honda, Y., Kim, H., Hong, Y.-C. (2015). Non-Linear Concentration-Response Relationships between Ambient Ozone and Daily Mortality. *Plos One, 10*(6). doi: 10.1371/journal.pone.0129423

Bass, V., Gordon, C., Jarema, K., Macphail, R., Cascio, W., Phillips, P., Ledbetter, A., Schladweiler, M., Andrews, D., Miller, D., Doerfler, D., Kodavanti, U. (2013). Ozone induces glucose intolerance and systemic metabolic effects in young and aged brown Norway rats. *Toxicology and Applied Pharmacology, 273*(3), 551–560. doi: 10.1016/j.taap.2013.09.029

Bergmann, E., Bender, J., Weigel, H.-J. (2017). Impact of tropospheric ozone on terrestrial biodiversity: A literature analysis to identify ozone sensitive taxa. *Journal of Applied Botany and Food Quality, 90,* 83-105. doi:10.5073/JABFQ.2017.090.012

Blanchard, C. L., Hidy, G. M. (2018). Ozone response to emission reductions in the southeastern United States. *Atmospheric Chemistry and Physics*, 18(11), 8183–8202. doi:10.5194/acp-18-8183-2018

Blanchard, C. L., Shaw, S. L., Edgerton, E. S., Schwab, J. J. (2019). Emission influences on air pollutant concentrations in New York State: I. ozone. *Atmospheric Environment: X, 3*, 100033. doi: 10.1016/j.aeaoa.2019.100033

Blande, J. D., Holopainen, J. K., Li, T. (2010). Air pollution impedes plant-to-plant communication by volatiles. *Ecology Letters, 13*(9), 1172–1181. doi: 10.1111/j.1461-0248.2010.01510.x

Blande, J. D., Holopainen, J. K., Niinemets, Ü. (2014). Plant volatiles in polluted atmospheres: stress responses and signal degradation. *Plant, Cell & Environment, 37*(8), 1892–1904. doi: 10.1111/pce.12352

Brauer, M., Brumm, J., Vedal, S., Petkau, A. J. (2002). Exposure Misclassification and Threshold Concentrations in Time Series Analyses of Air Pollution Health Effects. *Risk Analysis*, *22*(6), 1183–1193. doi: 10.1111/1539-6924.00282

Carrothers, T. J., Evans, J. S. (2000). Assessing the Impact of Differential Measurement Error on Estimates of Fine Particle Mortality. *Journal of the Air & Waste Management Association*, *50*(1), 65–74. doi: 10.1080/10473289.2000.10463988

Cefalu, M., Dominici, F. (2014). Does Exposure Prediction Bias Health-Effect Estimation?: The relationship between confounding adjustment and exposure prediction. *Epidemiology*, *25*(4), 583–590. doi: 10.1097/ede.0000000000000099

Chapin, F. S., Matson, P., Mooney, H. (2002). *Principles of terrestrial ecosystem ecology*. New York: Springer.

Chen, B.-Y., Chen, C.-H., Chuang, Y.-C., Kim, H., Honda, Y., Chiang, H.-C., Guo, Y. L. (2016a). Schoolchildrens antioxidation genotypes are susceptible factors for reduced lung function and airway inflammation caused by air pollution. *Environmental Research*, *149*, 145–150. doi: 10.1016/j.envres.2016.05.007

Chen, K., Wolf, K., Hampel, R., Stafoggia, M., Breitner, S., Cyrys, J., Samoli, E., Andersen, Z. J., Bero-Bedada, G., Bellander, T., Hennig, F., Jacquemin, B., Pekkanen, J., Peters, A., Schneider, A. (2018). Does temperature-confounding control influence the modifying effect of air temperature in ozone–mortality associations? *Environmental Epidemiology*, *2*(1). doi: 10.1097/ee9.0000000000000008

Chen, Z., Salam, M. T., Toledo-Corral, C., Watanabe, R. M., Xiang, A. H., Buchanan, T. A., Habre, R., Bastain, T. M., Lurmann, F., Wilson, J. P., Trigo, E., Gilliland, F. D. (2016b). Ambient Air Pollutants Have Adverse Effects on Insulin and Glucose Homeostasis in Mexican Americans. *Diabetes Care*, *39*(4), 547–554. doi: 10.2337/dc15-1795

Chen, Z., Wang, X.-K., Yao, F.-F., Zheng, F.-X., Feng, Z.-Z. (2010). Elevated Ozone Changed Soil Microbial Community in a Rice Paddy. *Soil Science Society of America Journal*, *74*(3), 829. doi: 10.2136/sssaj2009.0258

Chieppa, J., Chappelka, A., Eckhardt, L. (2015). Effects of tropospheric ozone on loblolly pine seedlings inoculated with root infecting ophiostomatoid fungi. *Environmental Pollution*, *207*, 130–137. doi: 10.1016/j.envpol.2015.08.053

Cornelissen, T. (2011). Climate change and its effects on terrestrial insects and herbivory patterns. *Neotropical Entomology*, *40*(2), 155–163. doi: 10.1590/s1519-566x2011000200001

Cox, L. A. (2018). Effects of exposure estimation errors on estimated exposure-response relations for PM2.5. *Environmental Research*, *164*, 636–646. doi: 10.1016/j.envres.2018.03.038

Cui, H., Wei, J., Su, J., Li, C., Ge, F. (2016). Elevated O3 increases volatile organic compounds via jasmonic acid pathway that promote the preference of parasitoid Encarsia formosa for tomato plants. *Plant Science*, *253*, 243–250. doi: 10.1016/j.plantsci.2016.09.019

Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J. D., Dominici, F. (2017). Association of Short-term Exposure to Air Pollution With Mortality in Older Adults. *Jama*, *318*(24), 2446. doi: 10.1001/jama.2017.17923

Dietze, M. C., Moorcroft, P. R. (2011). Tree mortality in the eastern and central United States: patterns and drivers. *Global Change Biology*, *17*(11), 3312–3326. doi: 10.1111/j.1365-2486.2011.02477.x

Ebelt, S. T., Wilson, W. E., Brauer, M. (2005). Exposure to Ambient and Nonambient Components of Particulate Matter. *Epidemiology*, *16*(3), 396–405. doi: 10.1097/01.ede.0000158918.57071.3e

Fares, S., Vargas, R., Detto, M., Goldstein, A. H., Karlik, J., Paoletti, E., Vitale, M. (2013). Tropospheric ozone reduces carbon assimilation in trees: estimates from analysis of continuous flux measurements. *Global Change Biology*, *19*(8), 2427–2443. doi: 10.1111/gcb.12222

Farré-Armengol, G., Peñuelas, J., Li, T., Yli-Pirilä, P., Filella, I., Llusia, J., Blande, J. D. (2015). Ozone degrades floral scent and reduces pollinator attraction to flowers. *New Phytologist*, *209*(1), 152–160. doi: 10.1111/nph.13620

Fewell, Z., Smith, G. D., Sterne, J. A. C. (2007). The Impact of Residual and Unmeasured Confounding in Epidemiologic Studies: A Simulation Study. *American Journal of Epidemiology*, *166*(6), 646–655. doi: 10.1093/aje/kwm165

Fiscus, E. L., Booker, F. L., Sadok, W., Burkey, K. O. (2012). Influence of atmospheric vapour pressure deficit on ozone responses of snap bean (Phaseolus vulgaris L.) genotypes. *Journal of Experimental Botany*, *63*(7), 2557–2564. doi: 10.1093/jxb/err443

Folinsbee, L. J., Mcdonnell, W. F., Horstman, D. H. (1988). Pulmonary Function and Symptom Responses after 6.6-Hour Exposure to 0.12 ppm Ozone with Moderate Exercise. *Journal of the Air Pollution Control Association*, *38*(1), 28–35. doi: 10.1080/08940630.1988.10466349

Frampton, M. W., Balmes, J. R., Bromberg, P. A., Stark, P., Arjomandi, M., Hazucha, M. J., Rich, D. Q., Hollenbeck-Pringle, D., Dagincourt, N., Alexis, N., Ganz, P., Zareba, W., Constantini, M.G. (2017). Multicenter Ozone Study in oldEr Subjects (MOSES): Part 1. Effects of Exposure to Low Concentrations of Ozone on Respiratory and Cardiovascular Outcomes. *Health Effects Institute Research Report*, No. 192.

Frampton, M. W., Morrow, P. E., Torres, A., Cox, C., Voter, K. Z., Utell, M. J. (1997). Ozone responsiveness in smokers and nonsmokers. *American Journal of Respiratory and Critical Care Medicine*, *155*(1), 116–121. doi: 10.1164/ajrccm.155.1.9001299

Frampton, M. W., Pietropaoli, A., Dentler, M., Chalupa, D., Little, E. L., Stewart, J., Frasier, L., Oakes, D., Wiltshire, J., Vora, R., Utell, M. J. (2015). Cardiovascular effects of ozone in healthy subjects with and without deletion of glutathione-S-transferase M1. *Inhalation Toxicology*, *27*(2), 113–119. doi: 10.3109/08958378.2014.996272

Fuentes, J. D., Roulston, T. A. H., Zenker, J. (2013). Ozone impedes the ability of a herbivore to find its host. *Environmental Research Letters*, *8*(1), 014048. doi: 10.1088/1748-9326/8/1/014048

Garcia, E., Berhane, K. T., Islam, T., Mcconnell, R., Urman, R., Chen, Z., Gilliland, F. D. (2019). Association of Changes in Air Quality With Incident Asthma in Children in California, 1993-2014. *Journal of the American Medical Association*, *321*(19), 1906. doi: 10.1001/jama.2019.5357

Gillespie, C., Stabler, D., Tallentire, E., Goumenaki, E., Barnes, J. (2015). Exposure to environmentally-relevant levels of ozone negatively influence pollen and fruit development. *Environmental Pollution, 206*, 494–501. doi: 10.1016/j.envpol.2015.08.003

Gilliland, N. J., Chappelka, A. H., Muntifering, R. B., Ditchkoff, S. S. (2015). Changes in southern Piedmont grassland community structure and nutritive quality with future climate scenarios of elevated tropospheric ozone and altered rainfall patterns. *Plant Biology*, *18*, 47–55. doi: 10.1111/plb.12324

Goldman, G. T., Mulholland, J. A., Russell, A. G., Strickland, M. J., Klein, M., Waller, L. A., Tolbert, P. E. (2011). Impact of exposure measurement error in air pollution epidemiology: effect of error type in time-series studies. *Environmental Health*, *10*(1). doi: 10.1186/1476-069x-10-61

Gordon, C. J., Jarema, K. A., Lehmann, J. R., Ledbetter, A. D., Schladweiler, M. C., Schmid, J. E., Ward, W. O., Kodavanti, U. P., Nyska, A., Macphail, R. C. (2013). Susceptibility of adult and senescent Brown Norway rats to repeated ozone exposure: an assessment of behavior, serum biochemistry and cardiopulmonary function. *Inhalation Toxicology*, *25*(3), 141–159. doi: 10.3109/08958378.2013.764946

Gordon, C. J., Phillips, P. M., Johnstone, A. F. M., Beasley, T. E., Ledbetter, A. D., Schladweiler, M. C., Snow, S.J., Kodavanti, U. P. (2016). Effect of high-fructose and high-fat diets on pulmonary sensitivity, motor activity, and body composition of brown Norway rats exposed to ozone. *Inhalation Toxicology, 28*(5), 203–215. doi: 10.3109/08958378.2015.1134730

Hansen, J. S., Nørgaard, A. W., Koponen, I. K., Sørli, J. B., Paidi, M. D., Hansen, S. W. K., Clausen, P. A., Nielsen, G. D., Wolkoff, P., Larsen, S. T. (2016). Limonene and its ozone-initiated reaction products attenuate allergic lung inflammation in mice. *Journal of Immunotoxicology*, *13*(6), 793–803. doi: 10.1080/1547691x.2016.1195462

Hart, J. E., Liao, X., Hong, B., Puett, R. C., Yanosky, J. D., Suh, H., Kioumourtzoglou, M.-A., Spiegelman, D., Laden, F. (2015). The association of long-term exposure to PM2.5 on all-cause mortality in the Nurses' Health Study and the impact of measurement-error correction. *Environmental Health, 14*(1). doi: 10.1186/s12940-015-0027-6

Hatch, G. E., Mckee, J., Brown, J., Mcdonnell, W., Seal, E., Soukup, J., Slade, R., Crissman, K., Devlin, R. (2013). Biomarkers of Dose and Effect of Inhaled Ozone in Resting versus Exercising Human Subjects: Comparison with Resting Rats. *Biomarker Insights*, *8*. doi: 10.4137/bmi.s11102

Hatch, G. E., Slade, R., Harris, L. P., Mcdonnell, W. F., Devlin, R. B., Koren, H. S., Costa, D. L., Mckee, J. (1994). Ozone dose and effect in humans and rats. A comparison using oxygen-18 labeling and bronchoalveolar lavage. *American Journal of Respiratory and Critical Care Medicine*, *150*(3), 676–683. doi: 10.1164/ajrccm.150.3.8087337

Holmes, W. E., Zak, D. R., Pregitzer, K. S., King, J. S. (2006). Elevated CO2 and O3 Alter Soil Nitrogen Transformations beneath Trembling Aspen, Paper Birch, and Sugar Maple. *Ecosystems*, *9*(8), 1354–1363. doi: 10.1007/s10021-006-0163-5

Jaffe, D. A., Cooper, O. R., Fiore, A. M., Henderson, B. H., Tonneson, G. S., Russell, A. G., Henze, D. K., Langford, A. O., Lin, M., Moore, T. (2018). Scientific assessment of background ozone over the U.S.: Implications for air quality management. *Elem Sci Anth*, *6*(1), 56. doi: 10.1525/elementa.309

Jerrett, M., Brook, R., White, L. F., Burnett, R. T., Yu, J., Su, J., Seto, E., Marshall, J., Palmer, J. R., Rosenberg, L., Coogan, P. F. (2017). Ambient ozone and incident diabetes: A prospective analysis in a large cohort of African American women. *Environment International*, *102*, 42–47. doi: 10.1016/j.envint.2016.12.011

Jurek, A. M., Greenland, S., Maldonado, G. (2008). Brief Report: How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? *International Journal of Epidemiology*, 37(2), 382–385. doi:10.1093/ije/dym291

Jurek, A. M., Greenland, S., Maldonado, G., Church, T. R. (2005). Proper interpretation of non-differential misclassification effects: expectations vs observations. *International Journal of Epidemiology*, *34*(3), 680–687. doi: 10.1093/ije/dyi060

Kanerva, T., Palojärvi, A., Rämö, K., Ojanperä, K., Esala, M., Manninen, S. (2006). A 3-year exposure to CO2 and O3 induced minor changes in soil N cycling in a meadow ecosystem. *Plant and Soil*, *286*(1-2), 61–73. doi: 10.1007/s11104-006-9026-2

Kendall, R. J., Lacher, T. E., Cobb, G. P., Cox, S. B. (Eds.). (2010). *Wildlife Toxicology: Emerging Contaminant and Biodiversity Issues*. S.l.: CRC Press/Taylor and Francis Group.

Kohut, R. (2007). Assessing the risk of foliar injury from ozone on vegetation in parks in the U.S. National Park Services Vital Signs Network. *Environmental Pollution*, *149*(3), 348–357. doi: 10.1016/j.envpol.2007.04.022

Körner, C. (2006). Plant CO2responses: an issue of definition, time and resource supply. *New Phytologist*, *172*(3), 393–411. doi: 10.1111/j.1469-8137.2006.01886.x

Leisner, C. P., Ainsworth, E. A. (2011). Quantifying the effects of ozone on plant reproductive growth and development. *Global Change Biology*, *18*(2), 606–616. doi: 10.1111/j.1365-2486.2011.02535.x

Leisner, C. P., Yendrek, C. R., Ainsworth, E. A. (2017). Physiological and transcriptomic responses in the seed coat of field-grown soybean (Glycine max L. Merr.) to abiotic stress. *BMC Plant Biology*, *17*(1). doi: 10.1186/s12870-017-1188-y

Li, T., Blande, J. D., Holopainen, J. K. (2016). Atmospheric transformation of plant volatiles disrupts host plant finding. *Scientific Reports*, *6*(1). doi: 10.1038/srep33851

Lipfert, F.W., Wyzga, R., (1996). The effects of exposure error on environmental epidemiology. In: Proceedings of the Second Colloquium on Particulate Air Pollution and Human Mortality and Morbidity. pp. 295–302.

McCant, D., Lange, S., Haney, J., Honeycutt, M. (2017). The perpetuation of the misconception that rats receive a 3–5 times lower lung tissue dose than humans at the same ozone concentration. *Inhalation Toxicology*, *29*(5), 187–196. doi: 10.1080/08958378.2017.1323982

McDonnell, W. F., Kehrl, H. R., Abdul-Salaam, S., Ives, P. J., Folinsbee, L. J., Devlin, R. B., O'Neil, J. J., Horstman, D. H. (1991). Respiratory Response of Humans Exposed to Low Levels of Ozone for 6.6 Hours. *Archives of Environmental Health: An International Journal*, *46*(3), 145–150. doi: 10.1080/00039896.1991.9937441

McGrath, J. M., Betzelberger, A. M., Wang, S., Shook, E., Zhu, X.-G., Long, S. P., Ainsworth, E. A. (2015). An analysis of ozone damage to historical maize and soybean yields in the United States. *Proceedings of the National Academy of Sciences*, *112*(46), 14390–14395. doi: 10.1073/pnas.1509777112

McGuinn, L. A., Ward-Caviness, C., Neas, L. M., Schneider, A., Di, Q., Chudnovsky, A., Schwartz, J., Koutrakis, P., Russell, A. G., Garcia, V., Kraus, W. E., Hauser, E. R., Cascio, W., Diaz-Sanchez, D., Devlin, R. B. (2017). Fine particulate matter and cardiovascular disease: Comparison of assessment methods for long-term exposure. *Environmental Research*, *159*, 16–23. doi: 10.1016/j.envres.2017.07.041

Miller, D. B., Ghio, A. J., Karoly, E. D., Bell, L. N., Snow, S. J., Madden, M. C., Soukup, J., Cascio, W. E., Gilmour, M. I.,, Kodavanti, U. P. (2016a). Ozone Exposure Increases Circulating Stress Hormones and Lipid Metabolites in Humans. *American Journal of Respiratory and Critical Care Medicine*, *193*(12), 1382–1391. doi: 10.1164/rccm.201508-1599oc

Miller, D. B., Snow, S. J., Henriquez, A., Schladweiler, M. C., Ledbetter, A. D., Richards, J. E., Andrews, D. L., Kodavanti, U. P. (2016b). Systemic metabolic derangement, pulmonary effects, and insulin insufficiency following subchronic ozone exposure in rats. *Toxicology and Applied Pharmacology*, *306*, 47–57. doi: 10.1016/j.taap.2016.06.027

Moolgavkar, S. H., Mcclellan, R. O., Dewanji, A., Turim, J., Luebeck, E. G., Edwards, M. (2013). Time-Series Analyses of Air Pollution and Mortality in the United States: A Subsampling Approach. *Environmental Health Perspectives*, *121*(1), 73–78. doi: 10.1289/ehp.1104507

Moore, K., Neugebauer, R., Lurmann, F., Hall, J., Brajer, V., Alcorn, S., Tager, I. (2008). Ambient Ozone Concentrations Cause Increased Hospitalizations for Asthma in Children: An 18-Year Study in Southern California. *Environmental Health Perspectives*, *116*(8), 1063–1070. doi: 10.1289/ehp.10497

Moore, K. L., Neugebauer, R., Laan, M. J., Tager, I. B. (2012). Causal inference in epidemiological studies with strong confounding. *Statistics in Medicine*, *31*(13), 1380–1404. doi: 10.1002/sim.4469

Moran, E. V., Kubiske, M. E. (2013). Can elevated CO2and ozone shift the genetic composition of aspen (Populus tremuloides) stands? *New Phytologist*, *198*(2), 466–475. doi: 10.1111/nph.12153

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., Zhang, H. (2013). Anthropogenic and natural radiative forcing. In Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M. (Ed.). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 659-740). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. doi:10.1017/CBO9781107415324.018.

Parrish, D. D., Ennis, C. A. (2019). Estimating background contributions and US anthropogenic enhancements to maximum ozone concentrations in the northern US. *Atmospheric Chemistry and Physics*, *19*(19), 12587–12605. doi: 10.5194/acp-19-12587-2019

Parrish, D. D., Young, L. M., Newman, M. H., Aikin, K. C., Ryerson, T. B. (2017). Ozone Design Values in Southern Californias Air Basins: Temporal Evolution and U.S. Background Contribution. *Journal of Geophysical Research: Atmospheres*, *122*(20). doi: 10.1002/2016jd026329

Pleijel, H., Broberg, M. C., Uddling, J., Mills, G. (2018). Current surface ozone concentrations significantly decrease wheat growth, yield and quality. *Science of The Total Environment*, *613-614*, 687–692. doi: 10.1016/j.scitotenv.2017.09.111

Ramot, Y., Kodavanti, U. P., Kissling, G. E., Ledbetter, A. D., Nyska, A. (2015). Clinical and pathological manifestations of cardiovascular disease in rat models: the influence of acute ozone exposure. *Inhalation Toxicology*, *27*(sup1), 26–38. doi: 10.3109/08958378.2014.954168

Rhomberg, L. R., Chandalia, J. K., Long, C. M., Goodman, J. E. (2011). Measurement error in environmental epidemiology and the shape of exposure-response curves. *Critical Reviews in Toxicology*, *41*(8), 651–671. doi: 10.3109/10408444.2011.563420

Rich, D. Q., Balmes, J. R., Frampton, M. W., Zareba, W., Stark, P., Arjomandi, M., Hazucha, M. J., Costantini, M. G., Ganz, P., Hollenbeck-Pringle, D., Dagincourt, N., Bromberg, P. A. (2018). Cardiovascular function and ozone exposure: The Multicenter Ozone Study in oldEr Subjects (MOSES). *Environment International*, *119*, 193–202. doi: 10.1016/j.envint.2018.06.014

Sanz, J., González-Fernández, I., Elvira, S., Muntifering, R., Alonso, R., Bermejo-Bermejo, V. (2016). Setting ozone critical levels for annual Mediterranean pasture species: Combined analysis of open-top chamber experiments. *Science of The Total Environment*, *571*, 670–679. doi: 10.1016/j.scitotenv.2016.07.035

Schelegle, E. S., Morales, C. A., Walby, W. F., Marion, S., Allen, R. P. (2009). 6.6-Hour Inhalation of Ozone Concentrations from 60 to 87 Parts per Billion in Healthy Humans. *American Journal of Respiratory and Critical Care Medicine*, *180*(3), 265–272. doi: 10.1164/rccm.200809-1484oc

Seiler, L. K., Decoteau, D. R., Davis, D. D. (2014). Evaluation of Ailanthus altissimaas a Bioindicator to Detect Phytotoxic Levels of Ozone. *Northeastern Naturalist*, *21*(4), 541–553. doi: 10.1656/045.021.0405

Seltzer, K. M., Shindell, D. T., Malley, C. S. (2018). Measurement-based assessment of health burdens from long-term ozone exposure in the United States, Europe, and China. *Environmental Research Letters*, *13*(10), 104018. doi: 10.1088/1748-9326/aae29d

Silverman, R. A., Ito, K. (2010). Age-related association of fine particles and ozone with severe acute asthma in New York City. *Journal of Allergy and Clinical Immunology*, *125*(2). doi: 10.1016/j.jaci.2009.10.061

Slikker, W., Andersen, M., Bogdanffy, M., Bus, J., Cohen, S., Conolly, R., David, R., Doerrer, N., Dorman, D., Gaylor, D., Hattis, D., Rogers, J.M., Setzer, R.W., Swenberg, J.A., Wallace, K. (2004a). Dose-dependent transitions in mechanisms of toxicity: case studies. *Toxicology and Applied Pharmacology*, *201*(3), 226–294. doi: 10.1016/j.taap.2004.06.027

Slikker, W., Andersen, M., Bogdanffy, M., Bus, J., Cohen, S., Conolly, R., David, R.M., Doerrer, N.G., Dorman, D.C., Gaylor, D.W., Hattis, D., Rogers, J.M., Setzer, R.W., Swenberg, J.A., Wallace, K. (2004b). Dose-dependent transitions in mechanisms of toxicity. *Toxicology and Applied Pharmacology*, *201*(3), 203–225. doi: 10.1016/j.taap.2004.06.019

Smith, G.C., Coulston, J.W., O'Connell, B.M. (2008). Ozone bioindicators and forest health: a guide to the evaluation, analysis, and interpretation of the ozone injury data in the Forest Inventory and Analysis Program. (NRS-34). Newtown Square, PA: U.S. Department of Agriculture, Forest Service. doi:10.2737/NRS-GTR-34.

Sorahan, T., Gilthorpe, M. S. (1994). Non-differential misclassification of exposure always leads to an underestimate of risk: an incorrect conclusion. *Occupational and Environmental Medicine*, *51*(12), 839–840. doi: 10.1136/oem.51.12.839

Stoelken, G., Pritsch, K., Simon, J., Mueller, C. W., Grams, T. E. E., Esperschuetz, J., Gayler, S., Buegger, F., Brueggemann, N., Meier, R., Zeller, B., Winkler, J. B., Rennenberg, H. (2010). Enhanced ozone exposure of European beech (Fagus sylvatica) stimulates nitrogen mobilization from leaf litter and nitrogen accumulation in the soil. *Plant Biosystems - An International Journal Dealing with All Aspects of Plant Biology*, *144*(3), 537–546. doi: 10.1080/11263500903429346

Talhelm, A. F., Pregitzer, K. S., Kubiske, M. E., Zak, D. R., Campany, C. E., Burton, A. J., Dickson, R. E., Hendrey, G. R., Isebrands, J. G., Lewin, K. F., Nagy, J., Karnosky, D. F. (2014). Elevated carbon dioxide and ozone alter productivity and ecosystem carbon content in northern temperate forests. *Global Change Biology*, *20*(8), 2492–2504. doi: 10.1111/gcb.12564

Tétreault, L.-F., Doucet, M., Gamache, P., Fournier, M., Brand, A., Kosatsky, T., Smargiassi, A. (2016). Childhood Exposure to Ambient Air Pollutants and the Onset of Asthma: An Administrative Cohort Study in Québec. *Environmental Health Perspectives*, *124*(8), 1276–1282. doi: 10.1289/ehp.1509838

Trenga, C. A., Sullivan, J. H., Schildcrout, J. S., Shepherd, K. P., Shapiro, G. G., Liu, L.-J. S., Kaufman, J. D., Koenig, J. Q. (2006). Effect of Particulate Air Pollution on Lung Function in Adult and Pediatric Subjects in a Seattle Panel Study. *Chest*, *129*(6), 1614–1622. doi: 10.1378/chest.129.6.1614

Torres, A., Utell, M. J., Morow, P. E., Voter, K. Z., Whitin, J. C., Cox, C., Looney, R., Speers, D. M., Tsai, Y., Frampton, M. W. (1997). Airway Inflammation in Smokers and Nonsmokers with Varying Responsiveness to Ozone. *American Journal of Respiratory and Critical Care Medicine*, *156*(3), 728–736. doi: 10.1164/ajrccm.156.3.9601054

U.S. EPA. (2013). Integrated Science Assessment for Ozone and Related Photochemical Oxidants. EPA/600/R-10/076F. Research Triangle Park, NC. http://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=247492

U.S. EPA. (2015). Preamble to the Integrated Science Assessments. EPA/600/R-15/067. Research Triangle Park, NC. https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=310244

U.S. EPA. (2018). Integrated Science Assessment for Particulate Matter (External Review Draft). EPA/600/R-18/179. Research Triangle Park, NC. https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=341593

Valkama, E., Koricheva, J., Oksanen, E. (2007). Effects of elevated O3, alone and in combination with elevated CO2, on tree leaf chemistry and insect herbivore performance: a meta-analysis. *Global Change Biology*, *13*(1), 184–201. doi: 10.1111/j.1365-2486.2006.01284.x

Vitolo, C., Scutari, M., Ghalaieny, M., Tucker, A., Russell, A. (2018). Modeling Air Pollution, Climate, and Health Data Using Bayesian Networks: A Case Study of the English Regions. *Earth and Space Science*, *5*(4), 76–88. doi: 10.1002/2017ea000326

Voter, K.Z., Whitin, J.C., Torres, A., Morrow, P.E., Cox, C., Tsai, Y., Utell, M.J., Frampton, M.W. (2001). Ozone Exposure and the Production of Reactive Oxygen Species by Bronchoalveolar Cells in Humans. *Inhalation Toxicology*, *13*(6), 465–483. doi: 10.1080/08958370117715

Wang, B., Shugart, H. H., Shuman, J. K., Lerdau, M. T. (2016). Forests and ozone: productivity, carbon storage and feedbacks. *Scientific Reports*, *6*(1). doi: 10.1038/srep22133

Watt, M., Godden, D., Cherrie, J., Seaton, A. (1995). Individual exposure to particulate air pollution and its relevance to thresholds for health effects: a study of traffic wardens. *Occupational and Environmental Medicine*, *52*(12), 790–792. doi: 10.1136/oem.52.12.790

Wilson, A., Rappold, A. G., Neas, L. M., Reich, B. J. (2014). Modeling the effect of temperature on ozone-related mortality. *The Annals of Applied Statistics*, *8*(3), 1728–1749. doi: 10.1214/14-aoas754

Wittig, V. E., Ainsworth, E. A., Long, S. P. (2007). To what extent do current and projected increases in surface ozone affect photosynthesis and stomatal conductance of trees? A meta-analytic review of the last 3 decades of experiments. *Plant, Cell & Environment*, *30*(9), 1150–1162. doi: 10.1111/j.1365-3040.2007.01717.x

Wittig, V. E., Ainsworth, E. A., Naidu, S. L., Karnosky, D. F., Long, S. P. (2009). Quantifying the impact of current and future tropospheric ozone on tree biomass, growth, physiology and biochemistry: a quantitative meta-analysis. *Global Change Biology*, *15*(2), 396–424. doi: 10.1111/j.1365-2486.2008.01774.x

Ying, Z., Allen, K., Zhong, J., Chen, M., Williams, K. M., Wagner, J. G., Lewandowski, R., Sun, Q., Rajagopalan, S., Harkema, J. R. (2016). Subacute inhalation exposure to ozone induces systemic inflammation but not insulin resistance in a diabetic mouse model. *Inhalation Toxicology*, *28*(4), 155–163. doi: 10.3109/08958378.2016.1146808

Yoshimura, I. (1990). The effect of measurement error on the dose-response curve. *Environmental Health Perspectives*, *87*, 173–178. doi: 10.1289/ehp.9087173

Yuan, J. S., Himanen, S. J., Holopainen, J. K., Chen, F., Stewart, C. N. (2009). Smelling global climate change: mitigation of function for plant volatile organic compounds. *Trends in Ecology & Evolution*, *24*(6), 323–331. doi: 10.1016/j.tree.2009.01.012

Yue, X., Unger, N. (2014). Ozone vegetation damage effects on gross primary productivity in the United States. *Atmospheric Chemistry and Physics*, *14*(17), 9137–9153. doi: 10.5194/acp-14-9137-2014

Zaccai, J. H. (2004). How to assess epidemiological studies. *Postgraduate Medical Journal*, *80*(941), 140–147. doi: 10.1136/pgmj.2003.012633

Zak, D. R., Kubiske, M. E., Pregitzer, K. S., Burton, A. J. (2011). Atmospheric CO2 and O3 alter competition for soil nitrogen in developing forests. *Global Change Biology*, *18*(4), 1480–1488. doi: 10.1111/j.1365-2486.2011.02596.x

Zhong, J., Allen, K., Rao, X., Ying, Z., Braunstein, Z., Kankanala, S. R., Xia, C., Wang, X., Bramble, L. A., Wagner, J. G., Lewandowski, R., Sun, Q., Harkema, J. R., Rajagopalan, S. (2016). Repeated ozone exposure exacerbates insulin resistance and activates innate immune response in genetically susceptible mice. *Inhalation Toxicology*, *28*(9), 383–392. doi: 10.1080/08958378.2016.1179373

Appendix A

Individual Comments by CASAC Members on the EPA's
Integrated Science Assessment for Ozone and Related Photochemical Oxidants
(External Review Draft – September 2019)

Dr. James Boylan

Executive Summary

Figure ES-2 on page ES-6 should change the asterisks (*) to up or down arrows to show upgraded and downgraded classifications.

Integrated Synthesis

Figure IS-6 on page IS-84 should change the asterisks (*) to up or down arrows to show upgraded and downgraded classifications.

Appendix 1 – Atmospheric Source, Chemistry, Meteorology, Trends, and Background

To what extent is the information presented in Appendix 1 regarding sources, precursor emissions, and measurement and modeling of ambient concentrations, as well as modeled estimates of background concentrations of ozone, clearly and accurately conveyed and appropriately characterized? Please comment on the extent to which available information on the spatial and temporal trends of ozone concentrations at various scales has been adequately and accurately described.

Sources of U.S. Ozone and its Precursors (Section 1.3)

This section presents estimated national values for 2014/2017 NEI emissions. However, there is no detailed discussion on the uncertainty associated with each pollutant or source sector. Some pollutants and sectors will be much more uncertain than others. For example, NOx emissions from electric generating units (EGUs) have low uncertainty since they are typically measured by hourly CEMs. On the other hand, other source sectors and pollutants may be highly uncertain. The uncertainties in the emissions inventory (magnitude, spatial allocation, and temporal allocation) should be discussed for each pollutant and source sector. In addition, it would be helpful to add national maps containing county-level emissions for NOx, VOCs, CO, and CH_4 to show the variability across the country.

It is not clear if CH_4 is included in the VOC emissions or not. The text should clearly state if CH_4 is included or excluded from the VOC emissions discussed in this Appendix. Due to the importance of biogenic VOCs, this section should discuss the differences between the BEIS and MEGAN models that are typically used to estimate biogenic VOC emissions. In addition, biogenic VOC trends should be included to see the variability from year-to-year and season-to-season.

Ozone Photochemistry (Section 1.4)

This section should start with a discussion of why the precursor emissions discussed in Section 1.3 (NOx, VOCs, CO, and CH_4) are important for ozone formation. An overview of the chemical mechanism should be presented, and important chemical reactions should be highlighted. The relative importance of each precursor should be discussed relative to urban ozone formation vs. USB ozone formation.

Inter-Annual Variability and Longer Term Trends in Meteorological Effects on Anthropogenic and U.S. Background Ozone (Section 1.5)

This section should discuss the impact of inter-annual variability and longer term trends in meteorological effects on ozone design values.

Measurements and Modeling (Section 1.6)

Ground-based ozone lidar instruments measure the vertical structure of ozone and quantify the mixing of plumes aloft. A review of these instruments and their capability should be added to this section. The section on "Satellite-Based Remote Sensing Methods" should include a discussion of the new TROPOMI satellite data that includes high resolution measurements of NO2 and formaldehyde. The section on "Advances in Regional Chemical Transport Modeling" should discuss the importance of performing a comprehensive model performance evaluation when using regional chemical transport models. This evaluation should include an evaluation of precursor pollutants to help ensure the model does not have compensating errors.

EPA's 2016 Exceptional Events Rule allows certain ozone measurements due to natural events to be excluded from the official design values when compared to the NAAQS. In some cases, identical exceptional events can be treated differently in one location vs. another based on how close the area is to the standard. In both locations, people are impacted by adverse health effects, but the data is removed in one location and not the other. The ISA should discuss how exceptional events are accounted for in health studies and risk analyses.

Ambient Air Concentrations and Trends (Section 1.7)

This section should discuss the shifting of ozone peak concentrations from summer to spring and fall that is occurring in many parts of the country (Blanchard and Hidy, 2018; Blanchard et al., 2019). In addition, this section should include a discussion on ozone precursor trends in addition to ozone trends. Specifically, trends in NOx, VOCs, and CO measurements form national monitoring networks (AQS, near-road, NCore, and PAMS) should be included and discussed.

U.S. Background Ozone Concentrations (Section 1.8)

Section 1.8.1 begins with the statement "As described in Section 1.2.2.1, USB ozone cannot be reliably estimated using ambient monitoring data because monitors can be influenced by U.S. emissions, including both relatively nearby emissions and interstate and hemispheric transport of ozone produced from U.S. emissions." Parrish et al. (2017) and Parrish and Ennis (2019) have shown that USB ozone can be reliably estimated using ambient monitoring data. Although monitors can be influenced by U.S. emissions, it is possible to account for these influences. Estimates from measurement-based approaches and from modeling-based approaches can be compared to understand differences and minimize the uncertainty in USB ozone estimates.

Emission controls have reduced ozone in the U.S. to the extent that background ozone contributes the majority of urban ozone concentrations, even on many days when ozone exceeds the NAAQS. Figures 1 and 2 show estimates of the ozone design values that would be present in the absence of U.S. or North

American anthropogenic emissions. Figure 1 is from a model calculation using the "zero-out sensitivity approach" (Jaffe et al., 2018). Figure 2 is developed from an observational-based approach (Parrish et al., 2017; Parrish and Ennis, 2019) applied to the entire country. These two maps show that in the southwestern U.S., background ozone makes such a large contribution that it will be extremely difficult to reach the 70 ppb NAAQS unless the background contribution decreases. Section 1.8.2.1 discusses new USB and North American Background estimates, but all of these estimates are for seasonal means. It is critical to evaluate the ozone design values that can result from USB.

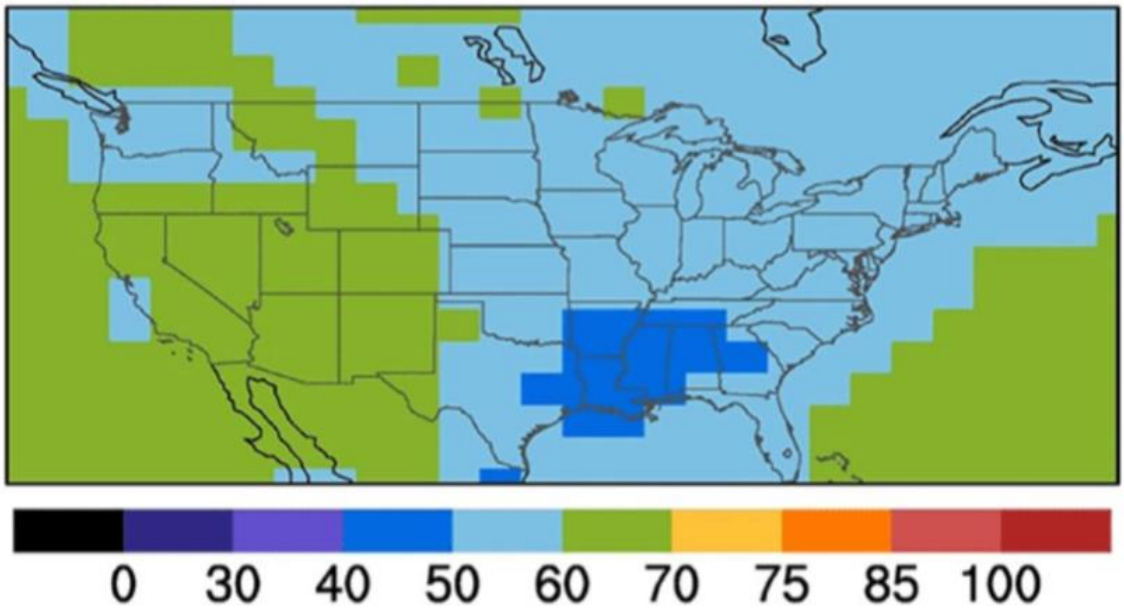


Figure 1. Annual 4th highest MDA8 O3 in ppb from North American background (i.e., with North American anthropogenic precursor emissions set to zero) averaged over 2010–2014 from a GFDL-AM3 model simulation (Jaffe et al., 2018).

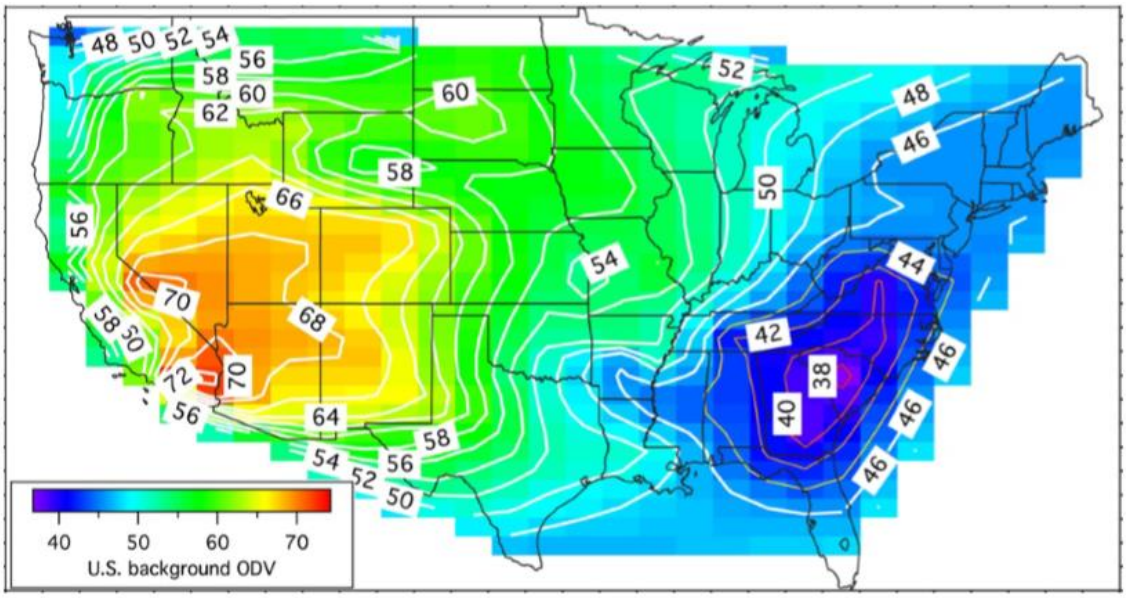


Figure 2. Ozone design values expected from U.S. background (i.e., with U.S. anthropogenic precursor emissions set to zero) in ~ 2015 derived from observations (D.D. Parrish, unpublished figure).

Appendix 2 – Exposure to Ambient Ozone

Appendix 2 describes scientific information on exposure to ozone and implications for epidemiologic studies. To what extent is the discussion on methodological considerations for exposure measurement and modeling clearly and accurately conveyed and appropriately characterized? Please comment on the extent to which the discussion regarding exposure assessment and the influence of exposure error on effect estimates in epidemiologic studies of the health effects of ozone has been adequately and accurately described.

Exposure Assessment Methods (Section 2.3)

This section gives a high-level overview of fixed-site monitors, passive and active personal samplers, spatial interpolation, land use regression and spatiotemporal modeling, chemical transport modeling, hybrid approaches, and microenvironmental modeling. The discussion on microenvironmental modeling should include additional information on APEX and SHEDS models.

Personal Exposure (Section 2.4)

This section discusses updates to the Consolidated Human Activity Database (CHAD), infiltration of ambient ozone into homes and buildings (I/O ratio), and personal exposure to ambient concentration (P/A) ratios. Additional discussion should be added for ozone infiltration in vehicles since a large amount of time is spent commuting. Also, a detailed discussion of the uncertainties and variability associated with the CHAD, I/O ratios, and P/A ratios should be included.

Copollutant Correlations and Potential for Confounding (Section 2.5)

It is stated on page 2-32 "Given that the majority of the copollutant correlation data are low, confounding of the relationship between ambient ozone exposure and a health effect by exposure to CO, SO_2, NO_2, PM_{10}, or $PM_{2.5}$ is less of a concern for studies of the health effects of ambient ozone exposure compared with studies of the health effects related to exposure of other criteria air pollutants. When copollutant correlations are higher during the warm season, greater risk of copollutant confounding exists." However, the summer is the season with the highest ozone concentrations and the highest ozone exposure; therefore, a greater risk of copollutant confounding exists and should be accounted for in the epidemiological studies.

Interpreting Exposure Measurement Error for Use in Epidemiology Studies (Section 2.6)

The summary table provided in the EPA ozone ISA presentation to CASAC on December 4, 2019 showing the influence of exposure error on epidemiology study outcomes (page 18) is a very useful summary and should be included in Appendix 2.

References

Blanchard, C. L. and G. M. Hidy, "Ozone response to emission reductions in the southeastern United States," Atmos. Chem. And Phys. 18: 8183–8202 (2018).

Blanchard, C. L., et.al., "Emission influences on air pollutant concentrations in New York State: I. Ozone," Atmospheric Environment: X 3 (2019) 100033

Jaffe D. A., et al. (2018) Scientific assessment of background ozone over the U.S.: Implications for air quality management. Elem. Sci. Anth., 6 56 doi.org/10.1525/elementa.309.

Parrish, D. D., Young, L. M., Newman, M. H., Aikin, K. C., and Ryerson, T. B. (2017) Ozone Design Values in Southern California's Air Basins: Temporal Evolution and U.S. Background Contribution, J. Geophys. Res.-Atmos., 122, 11166–11182, https://doi.org/10.1002/2016JD026329.

Parrish, D. D. and C. A. Ennis (2019). Estimating background contributions and US anthropogenic enhancements to maximum ozone concentrations in the northern US, Atmos. Chem. Phys., 19, 12587–12605, https://doi.org/10.5194/acp-19-12587-2019.

Dr. Tony Cox

Responses to Charge Questions

The Executive Summary is intended to provide a concise synopsis of the key findings and conclusions of the Ozone ISA for a broad range of audiences. ***Please comment on the clarity*** *with which the Executive Summary communicates the key information from the draft ISA.* ***Please provide recommendations on information that should be added*** *or information that should be left for discussion in the Integrated Synthesis and accompanying appendices of the draft ISA.*

The key information provided in the draft ISA and its Executive Summary is unclear. Concerns about lack of clarity in how key results are derived, expressed, and communicated have been raised in numerous public comments in this and previous NAAQS review cycles. They are still not addressed in the current draft ISA.

In addition to inviting public comments, this review cycle for the first time gave a panel of external expert consultants an opportunity to comment directly on the following question: *Is the scientific information provided by the ISA clear?* Appendix B provides responses from the consultants to this and other questions. Their main answer to this question is no, for at least the following reasons:

- *Criteria for selecting and weighting studies, and how key conclusions are derived from them, are not clear.*
- *It is unclear how, if at all, conclusions would change if consistent criteria were systematically applied* for selecting, evaluating, summarizing, and synthesizing studies.
- *The draft ISA and its Executive Summary do not provide comprehensive quantitative uncertainty and sensitivity analyses* showing how conclusions change for plausible variations in assumptions, interpretations of undefined and vague terms, selection and weighting of studies, and judgments on which the conclusions depend.
- *Causal determination judgments appear to be ambiguous, subjective, and sometimes arbitrary.* Several external consultants commented that different people might well make the determinations in very different ways from the same data. The evidence presented often does not seem to clearly support one causal determination to the exclusion of others. These experts noted that they could not guess, for any particular body of evidence, which causal determination category EPA will choose to describe it. By this criterion, the causal determinations do not seem to follow clearly from the evidence presented, but incorporate an arbitrary (unpredictable) element. The draft ISA provides no clear objective basis for determining or predicting from facts and data which causal determination (if any) is right.
- *What the causal determinations mean, and what they imply for empirical observations, is unclear*. For example multiple external experts agreed that the term "causal" is used in the draft ISA without distinguishing among importantly different causal concepts. No distinction is made among necessary, sufficient, INUS, and other forms of causation. Yet, a stated conclusion such as a determination that a specific C-R association is "causal" or "likely to be causal," is often

correct for some of these causal concepts and incorrect for others. Thus, the use of vague general terms such as "causal" or "likely to be causal" to communicate key conclusions in the draft ISA makes their meanings highly ambiguous and impossible to determine with clarity. As stated by Dr. Rhomberg, "If patterns of association are plausibly explained by underlying causation, this is taken [in the draft ISA] as sufficient evidence for such causation (when one should actually be comparing the hypothesized causative actions against other competing explanations for the patterns), and if such causation is inferred, it is taken to be universal, applying to other settings on the usually poorly stated (much less justified) presumption that the causation is universal and largely independent of other circumstances." Such inferences and presumptions, which pervade the draft ISA, are scientifically unsound, making the validity as well as the meanings of key conclusions unclear. Leaving unspecified exactly what the draft ISA means (e.g., one of these specific concepts, or perhaps something else) when it uses the term "causal" therefore makes it impossible to determine whether its key conclusions expressed using this term are correct, or even what they are intended to mean. This situation is sometimes referred to in other areas of science as "not even wrong," i.e., key findings depend on arbitrary-seeming judgments expressed using poorly defined, vague, or ambiguous terms, so that it is not clear what they mean (or how they could be tested or falsified by data), let alone whether they are correct.

- *The draft ISA's treatment of wildfire contributions to ozone exposure,* and their implications for NAAQS, are unclear.
- *It is unclear to what extent the ozone-associated physiological effects discussed in the draft ISA are transient* and to what extent they represent persistent, avoidable harms that could be reduced by further reducing ozone levels. Yet, this is crucial information for policy makers.

Information that should be added to the Executive Summary includes the following:

1. *Discussion of how changes in public health effects depend on changes in ozone levels.* This is the most important scientific topic for informing the PA. It is not addressed in the draft ISA. A useful ISA should address the extent to which reducing ozone has been found to cause reductions in public health risks and improvements in public health and welfare, and the extent to which additional reduction should or should not be expected to cause further benefits. It should quantify uncertainties about the answers. As stated by Dr. North, we "should be seeking to evaluate manipulative or interventional causation, that is, determining how many people might be added or subtracted from having their health protected with an adequate margin of safety by a change in the primary NAAQS standard." The draft ISA does not present relevant scientific information to use in addressing this question. Dr. Lipfert notes that "The ultimate test of causality is whether health has actually improved since the late 1970s in response to peak O3 levels reduced by a factor of 5 in conjunction with coincident trends in spatial patterns of reduced smoking and improved medical care. A search of PubMed found no support for such improvement." Rather than relying on searches by external experts, it would be far preferable for the EPA itself to address in the final ISA the key policy-relevant scientific question of how changes in public health risks depend on changes in ozone levels.

2. *Summary of results from a systematic review and critical evaluation and synthesis of relevant studies, including negative ones that have been omitted from the draft ISA.* The comments in Appendix B and other comments received from the external expert consultants and the public identify some of the omitted relevant studies (e.g., Moore et al. 2012).
3. *Detailed discussion of possible confounding*, and how it was or was not addressed in reaching causal conclusions. As noted by Dr. North in the context of answering a specific question, "I perceive that the kind of discussion needed on confounding was not present in the ISA, just a judgment of 'likely to be causal.' And I do not find such judgments useful in the absence of detailed discussion of possible confounding."
4. *Results of systematic evaluations of study quality, using consistently applied criteria,* showing how each key study included performs on each specific quality criterion relevant for evaluating individual studies and drawing valid causal conclusions (e.g., identification of appropriate adjustment sets, control of observed confounders, control of residual confounding, control of latent confounding, quantification of exposure estimation errors and uncertainties, adjustment of effects estimates for errors-in-variables, quantification of model uncertainty, adjustment of reported significance levels and confidence levels for model uncertainty, control for multiple testing bias, use of appropriate control groups, tests for internal validity, tests for external validity and invariant causal prediction property). A matrix (possibly color-coded, as suggested by Dr. Goodman in public comments on PM2.5) summarizing these results could provide great insight into the state of the literature and the strengths and limitations of individual key studies used by the EPA in reaching its conclusions.
5. *Discussion of causal biological mechanisms of inflammation-related health effects* and their implications for biologically realistic causal C-R functions. For example, the ISA should discuss recent evidence on mechanisms and modes of action (possibly including the role of the NLRP3 inflammasome in inflammation-mediated responses to ozone exposures) and their implications for the shapes of causal C-R functions describing health responses to changes in exposures below the current NAAQS. (Importantly, such causal C-R functions should not be confused with regression C-R functions.)
6. *Results of comprehensive, quantitative uncertainty and sensitivity analyses* showing how conclusions change for variations in inputs, including selection and weighting of studies, alternative interpretations of study results, corrections for confounding, corrections for measurement errors, corrections for historical trends, modeling choices and assumptions, interpretations of undefined and vague terms, and subjective judgments and unverified assumptions on which conclusions depend.

The Integrated Synthesis presents and synthesizes the overall conclusions from the subsequent detailed appendices of the draft ISA and characterizes available scientific information on policy relevant issues. Please comment on the ***usefulness and effectiveness of the summary*** *presentation. Please provide* ***recommendations on approaches that may improve the communication of key findings*** *to varied audiences and the* ***synthesis of available information*** *across subject areas.* ***What information should be added*** *or is more appropriate to leave for discussion in the subsequent detailed appendices?*

The Integrated Synthesis has the following limitations that limit its usefulness and effectiveness.

- *Biased selection of studies.* Multiple readers of the draft ISA, including several external expert consultants, have noted that it omits many relevant studies, especially those that do not conclude that ozone is associated with adverse health effects. This undermines the credibility, completeness, and scientific usefulness and effectiveness of the draft ISA, especially to the extent that it creates an impression that the evidence presented has been selected to support a narrative rather than to neutrally convey the current state of the art in the underlying science.
- *Literature on nonlinear effects is not well covered.* The draft ISA does not adequately cover the recent scientific literature on nonlinear C-R functions for ozone. For example, p. IS-88 states that "Examination of the concentration-response (C-R) relationship has primarily been conducted in studies of short-term ozone exposure and respiratory health effects or mortality, with some more recent studies characterizing this relationship for long-term ozone exposure and mortality. Across recent studies that used a variety of statistical methods to examine potential deviations from linearity, **evidence continues to support a linear C-R relationship**, but with less certainty in the shape of the curve at lower concentrations (i.e., below 30−40 ppb)." This contrasts with a substantial literature, disregarded in the draft ISA, on nonlinear C-R relationships. For example,
    - Bae et al. (2015) report that "The mean O3 concentration did not differ greatly between Korea and Japan, which were 26.2 ppb and 24.2 ppb, respectively. Seven out of 13 cities showed better fits for the spline model compared with the linear model, supporting a **non-linear relationships between O3 concentration and mortality**. **All of the 7 cities showed J or U shaped associations suggesting the existence of thresholds**. The range of city-specific thresholds was from 11 to 34 ppb. **The city-combined analysis also showed a non-linear association with a threshold around 30-40 ppb**." (Bae S, Lim YH, Kashima S, Yorifuji T, Honda Y, Kim H, Hong YC. Non-Linear Concentration-Response Relationships between Ambient Ozone and Daily Mortality.) PLoS One. 2015 Jun 15;10(6):e0129423. doi: 10.1371/journal.pone.0129423.
    - Seltzer et al. (2018) state that "Long-term ozone (O3) exposure estimates from chemical transport models are frequently paired with exposure-response relationships from epidemiological studies to estimate associated health burdens. **Impact estimates using such methods can include biases from model-derived exposure estimates**. We use data solely from dense ground-based monitoring networks in the United States, Europe, and China for 2015 to estimate long-term O3 exposure and calculate premature respiratory mortality using exposure-response relationships derived from two separate analyses of the American Cancer Society Cancer Prevention Study-II(ACS CPS-II) cohort. …**Both sets of results are lower (~20%–60%) on a region-by region basis than analogous prior studies based solely on modeled O3, due in large part to the fact that the latter tends to be high biased in estimating exposure.** This study highlights the utility of dense observation networks in estimating exposure to long-term O3 exposure and provides an observational constraint on subsequent health burdens for three regions of the world. In addition, **these results demonstrate how small biases in**

modeled results of long-term O3 exposure can amplify estimated health impacts due to nonlinear exposure-response curves."

- Wilson et al. (2014) report that, even in modeling that constrains ozone C-R functions for mortality to be monotonic (disallowing J-shaped or U-shaped relations such as those reported by Bae et al.), "We then examine the synergistic effect of ozone and temperature both nationally and locally and **find evidence of a nonlinear ozone effect** and an ozone-temperature interaction at higher temperatures and ozone concentrations." The draft ISA reports the nonlinear interaction from this study (p. 6-12), but does not mention the "evidence of a nonlinear ozone effect."

That the draft ISA does not mention such results on nonlinear C-R functions for ozone suggests that readers interested in understanding the available scientific evidence on ozone C-R functions must do their own research: the draft ISA fails to cover many studies and results that disagree with its narrative (in this case, that "evidence continues to support a linear C-R relationship"). This lack of coverage of diverse findings in the literature undermines the credibility, effectiveness and usefulness of the draft ISA.

- *Summaries of relevant literature are incomplete and of questionable accuracy.* The draft ISA does not provide a comprehensive or trustworthy summary of available scientific evidence, even for studies and authors that it cites. For example:
  - Page 3-91 of the ISA states that "A limited number of recent studies provide evidence of an association between long-term exposure to ozone and asthma development in children. … An overview of the evidence is provided below. A recent CHS analysis examined asthma incidence in relation to improved air quality in nine southern California communities (Garcia et al., 2019). ***Decreases in baseline ozone concentrations in three CHS cohorts, enrolled in 1993, 1996, and 2006, were associated with decreased asthma incidence.***" However, Garcia et al. (2019) actually state that "Among children in Southern California, decreases in ambient nitrogen dioxide and PM2.5 between 1993 and 2014 were significantly associated with lower asthma incidence. ***There were no statistically significant associations for ozone*** or PM10." (Garcia E, Berhane KT, Islam T, McConnell R, Urman R, Chen Z, Gilliland FD. Association of Changes in Air Quality With Incident Asthma in Children in California, 1993-2014. JAMA. 2019 May 21;321(19):1906-1915. doi: 10.1001/jama.2019.5357. Emphasis added.)
  - Table 3-3 on "Summary of evidence for a likely to be causal relationship between long-term ozone exposure and respiratory effects" cites the study of Moore et al. (2008) ("Ambient ozone concentrations cause increased hospitalizations for asthma in children: An 18-year study in Southern California") as providing "key evidence" for the ISA's causal determination that there is "a likely to be causal relationship between long-term ozone exposure and respiratory effects." Specifically, Moore et al. is cited as providing "***Consistent evidence of an association between long-term ozone concentrations and hospital admissions*** and ED visits for asthma." Yet, follow-up work by Moore et al. (2012) noted methodological limitations of the 2008 paper (especially, that its results

may have resulted from incorrect untested modeling assumptions, rather than from information in the data) and provided and applied an improved methodology ("CMRIER" or "causal models for realistic individualized exposure rules"). A key result was that the previous significant effect of ozone was no longer found. (Moore et al. (2012) state that "The results from the original HRMSM analysis based on the continuous ozone variable estimated with the G-computation method resulted in an estimate of an increase of 1.44e-06 in the proportion of asthma-related hospital discharges for a one-unit increase in ozone. [This is the 2008 study cited in Table 3-3 of the ISA.] ***Unlike results from the HRMSM analysis with the continuous ozone variable, the CMRIER results are not significant.*** Note that the HRMSM analysis was based on G-computation estimation which ***artificially relies on untestable parametric modeling assumptions*** to estimate HRMSM parameters when the ETA assumption is violated. Thus, in this ozone study [the 2008 study cited by the ISA], ***significant results from the G-computation analysis may be a consequence of the approach taken and not solely based on the information in the data***." (Moore KL, Neugebauer R, van der Laan MJ, Tager IB. Causal inference in epidemiological studies with strong confounding. Stat Med. 2012 Jun 15;31(13):1380-404. doi: 10.1002/sim.4469.) This more recent paper is not mentioned in the ISA. The ISA cites the 2008 results as "key evidence" without noting that the authors subsequently revised them in the 2012 paper.

- Table 3-3 cites a study by Tétreault et al. as providing "Key Evidence" of "Cohort studies demonstrating an *association* with asthma development in children." The ISA then interprets this, without any detailed explanation, as "Evidence for a *likely to be causal* relationship between long-term ozone exposure and respiratory effects."(Emphases added.) Yet, in discussing potential confounding, Tétreault et al. state that "We present two confounder models in the results. The first was adjusted for sex and deprivation, whereas the second was adjusted for the same variables as well as the year of birth*."* The article does not mention temperature or other weather variables. (For background on the importance of confounding by temperature, see e.g., Chen et al. (2018), "Does temperature-confounding control influence the modifying effect of air temperature in ozone-mortality associations?" This article concludes that using a categorical variable (e.g., a season indicator) to control for temperature yields highly significant ozone effects at high temperatures, but also significant residual confounding; and that adjusting for (nonlinear) effects of temperatures "substantially reduced ozone effects at high temperatures and residual confounding.") Tétreault et al. also note their "**lack of information on risk factors at the individual level** (e.g. socioeconomic status and smoking). We attempted to control for these factors with adjustments of our models using ecological deprivation variables, which are imperfect and **may result in residual confounding**." (Emphasis added.) Tétreault et al. further caution that "First, **individual exposure was modeled and not measured** through the follow-up, so the quality of the associations depends on the quality of the exposure models. All associations reported in

this study were estimated according to the exposure at the centroid of the residential postal code. This assumes that children would stay at home all day. Because a large proportion of a child's day can be spent outside the home (e.g., at school), where exposure to air pollutants might differ, **misclassification bias may have been introduced in our study**. Additionally, summer average O3 levels were used to estimate annual averages. Because summer O3 levels are higher than winter levels (Environment Canada 1999) in Canada, **we may have overestimated annual average levels**. Furthermore, although postal codes circumscribe a relatively small area in urban regions, postal codes may include much larger areas in rural regions. This difference in postal code size could lead to a degree of **higher imprecision in exposure estimation** in regions of the province that are less densely populated." (Emphasis added.) The ISA does not emphasize that the exposure concentrations that it reports (e.g., "32.1 ppb mean summer ozone concentration, based on 8-h midday avg" in Table 3-3) are in fact "modeled and not measured" values and does not adjust (e.g., using appropriate errors-in-variables methods) for potential biases due to such errors. It interprets the reported association as "key evidence of a likely causal relationship" without mentioning alternative interpretations such as that it might reflect omitted confounders (e.g., temperature), residual confounding, or misclassification bias. Page 3-193 of the ISA states that "Sensitivity analyses with alternate specifications for potential confounding inform the stability of findings and aid in judgments of the strength of inference from results." But it is not clear how or whether the ISA considered the results of such sensitivity analyses for the individual studies it relies on for its conclusions (e.g., in interpreting the Tétreault et al. study as "Key Evidence" of a "likely to be causal" relationship) or how sensitive the resulting causal determinations are to incompletely controlled confounding.

These examples are intended to be illustrative rather than comprehensive. They illustrate a larger issue: such coverage suggests that the draft ISA cannot serve as a trustworthy source for an accurate, unbiased, comprehensive critical summary and synthesis of the relevant scientific literature. As noted by multiple external experts and in public comments, the draft ISA appears to be biased toward defending EPA's methods and conclusions rather than providing a neutral, accurate review and summary and critical analysis and synthesis of available scientific studies. Its appearance of cherry-picking and bias in reporting results from the scientific literature undermines the effectiveness, trustworthiness, and usefulness of the draft ISA.

- *Policy-relevant science is not addressed.* The draft ISA does not usefully summarize, or critically evaluate, available scientific information on manipulative causation (i.e., on whether or to what extent reducing ozone reduces public health risks). Yet, this is the main topic needed to inform policy decisions about the public health consequences of alternative possible policy choices. For example, the external experts were directly asked "*Can valid determinations of manipulative or interventional causation – that is, how and whether changing exposure would change health risks – be made based on observed associations of the types analyzed in the ISA?*"

Most who answered said no; none said yes (see responses in Appendix B). For example, Dr. North stated that "I think this is a clear **NO**. CASAC should be seeking to evaluate manipulative or interventional causation, that is, determining how many people might be added or subtracted from having their health protected with an adequate margin of safety by a change in the primary NAAQS standard." Unless this omission is fixed, the PA lack a scientific foundation in the ISA for predicting effects on public health of alternative policies.

- *Uncertain relevance of facts addressed.* The draft ISA identifies several associations between ozone and physiological changes in controlled human experiments and epidemiological data, but it does not adequately address the extent to which these associations predict adverse effects on public health. Mr. Jansen frames the issue as follows: "In addition to the issue of beneficial effects, there is the issue of recovery or reversibility. …I did not see how it affected weighting nor causality classification. In other words, if a metric was responsive but recovered, how is that evidence weighted and used in terms of causality classification?" Similarly, Dr. North states that "It seems to me an important issue whether observed mild, apparently reversible effects such as changes in FEV1 (forced expiratory volume in one second) seen in healthy young exercising subjects imply a potential for adverse health effects in the general population. What are the adverse health effects, and how well do FEV1 changes predict them? What is the C-R relationship, not just for FEV1 changes, but for adverse health impacts that are persistent and perhaps cumulative over time, such as scarring of lung tissue so that lung function is permanently lost?" Health effects aside, the relevance of other information presented in the draft ISA is also often unclear. As stated by Dr. Parrish: "In reading through Section 1.3 a great deal of scientific information is summarized, but there is little or no discussion of the relevance of this science to the NAAQS or the ozone design values upon which the NAAQS is based." The final ISA should directly address the questions of the relevance of reversible effects, and of other information presented, to predicting public health responses to changes in ozone.

As mentioned above, the following additions to the draft ISA and Executive Summary are recommended to improve the communication of key results, and also the policy relevance, scientific validity, and methodological integrity of the content being communicated:

1. *Summarize available empirical evidence on how changes in public health effects depend on changes in ozone levels.*
2. *Present summary results from a systematic review and critical evaluation and synthesis of relevant studies,* including negative ones that have been omitted from the draft ISA.
3. *Provide detailed discussion of possible confounding*, and how it was or was not addressed for each study used to support causal conclusions.
4. *Present results of systematic evaluations of study quality*, using consistently applied criteria, showing how each key study included performs on each specific quality criterion relevant for drawing valid causal conclusions.
5. *Discuss causal biological mechanisms of inflammation-related health effects* preventable by reducing current ozone levels.

6. *Present comprehensive, quantitative uncertainty and sensitivity analyses* showing how the ISA's conclusions change for variations in selection and weighting of studies, modeling choices and assumptions, interpretations of undefined and vague terms, and subjective judgments on which the conclusions depend.

Additional Comments on Executive Summary

p. 1 "*This Integrated Science Assessment (ISA) is a* ***comprehensive evaluation and synthesis of the policy-relevant*** *science aimed at characterizing the health and welfare effects caused by ozone.*"

Comments:

- The draft ISA is not a comprehensive evaluation and synthesis. Multiple public comments and external expert consultant comments have pointed out that the draft ISA omits many relevant studies and topics; it appears to multiple reviewers to exhibit a selection bias favoring including positive studies while excluding negative ones, and does not address causal C-R functions for changes in public health risks caused by changes in ozone levels.
- The draft ISA does not include policy-relevant science, i.e., studies and empirical testing and validation of predictive generalizations that would allow changes in public health risks caused by alternative changes in NAAQA to be assessed, and uncertainties about them to be characterized.

p. 1 "***The ISA identifies and critically evaluates the most policy-relevant scientific literature across scientific disciplines****, including epidemiology, controlled human exposure studies, animal toxicology, atmospheric science, exposure science, vegetation studies, agricultural science, ecology, and climate-related science.* ***Key scientific conclusions (i.e., causality determinations; Section ES.4) are presented*** *and explained. They provide the scientific basis for developing risk and exposure analyses, policy evaluations, and policy decisions for the review. This ISA draws conclusions about the causal nature of the relationships between ozone exposure and health and welfare effects by integrating information across scientific disciplines and building off the evidence base evaluated in previous reviews.* ***The ISA thus provides the policy-relevant scientific information*** *that supports the review of the NAAQS.*"

Comments:

- *The highly relevant disciplines of health risk analysis, decision science, causal analysis, data science, and mathematical and simulation modeling are not adequately represented or used in the draft ISA.* For example, validation of health effect models is not discussed.

The causality determinations are not "key scientific conclusions." As discussed in more detail in several places in this document and in multiple public comments and external expert comments, the causal determinations are ambiguous expressions of subjective judgments. They do not provide a valid objective scientific basis for developing risk and exposure analyses, policy evaluations, or policy decisions. To be genuinely scientific, conclusions should rest on reproducible results of tests of unambiguously stated predictions against data. Neither this draft ISA nor the previous ISA for ozone presents conclusions that are "scientific" in this traditional sense. Rather, they apply the term "scientific" to ambiguously stated opinions and judgments. This should be fixed in the final ISA.

Dr. Mark Frampton

Limited expertise available to CASAC

The CASAC review of this ISA is limited by important changes in the review process that were recently implemented. For this ozone review, the EPA has failed to appoint an expert panel to assist CASAC in the review, as has been done for previous ozone reviews. The same panel of 12 consultants that was available for the PM review was available to respond to written questions from CASAC. However, notably this panel did not include any individuals actively participating in ozone health effects research, and did not include expertise in human clinical studies, which are critically important in understanding ozone health effects. These consultants did not attend the public meetings on the ISA, and there was no opportunity for interactive discussion. The limited expertise available for this review has adversely impacted CASAC's ability to provide the best advice to the Administrator.

Preamble

Study Quality. Section 4 (p. 7), regarding the assessment of study quality, does not indicate how quality assessments are used in the review. This issue was raised in the CASAC review of the PM ISA as well. The list of quality aspects that are reviewed are appropriate and complete, but nothing is provided about how these criteria are used or applied in the overall interpretation or assessment. It is not clear that the ISA consistently considers or incorporates these study quality assessments in reaching conclusions. This process should be strengthened and more fully described. The Preamble should provide details of how the study quality assessments are recorded, and of how they are considered in the development of the ISA and the PA.

Page 28. The publication referenced here, *What Constitutes an Adverse Health Effect of Air Pollution?* (ATS, 2000), should be updated with the latest version: *A joint ERS/ATS policy statement: what constitutes an adverse health effect of air pollution? An analytical framework* (Eur Respir J 2017). The statement on this page that this document "…described transient decrements in lung function as adverse when accompanied by clinical symptoms", while correct, over-simplifies the issue. Transient decrements in lung function should be considered adverse in some circumstances, even in the absence of symptoms. The older ATS document provided this statement as an example of <u>one</u> of the situations where transient decrements should be considered adverse.

Change in causality determination for short-term total mortality and cardiovascular effects: incomplete scientific review

Section 10.3.1.4. indicates, "In instances when a "causal" or "likely to be a causal" relationship was concluded in the 2013 Ozone ISA (i.e., short-term ozone exposure and respiratory and cardiovascular effects and total mortality, and long-term ozone exposure and respiratory effects), the epidemiologic studies evaluated for those outcomes were more limited in scope and targeted towards study locations that include U.S. airsheds or airsheds that are similar to those found in the U.S., as reflected in the PECOS tool."

The rationale for limiting epi studies in these categories of causality is to emphasize the studies most relevant for policy in addressing possible changes in the NAAQS. This is reasonable for outcomes

determined to be causal or likely to be causal. The problem is that, in the current ISA, for short-term total mortality and CV effects, the causality determinations were downgraded from likely to suggestive, based on the studies reviewed in the ISA, which were limited as indicated above. Part of the rationale for downgrading these causality determinations was continued limitations in the epidemiological evidence. We don't know from the ISA how many studies were excluded from consideration based on their location, or what was the impact (if any) of those exclusions the causality determination. The question is whether that causality determination would have been downgraded had all the evidence been considered. This needs to be addressed in the ISA, with a broadening of the epi review criteria, and re-assessment of the strength of the causality relationship, for these categories of health effects.

A brief PubMed search limited to the last 5 years identified more than 40 relevant epidemiology studies examining mortality and cardiovascular disease outcomes, conducted outside of North America. The following 3 studies appeared to be of particular high quality and relevance, and were published in high quality journals:

1. Bae S, Lim YH, Kashima S, Yorifuji T, Honda Y, Kim H, Hong YC. Non-Linear Concentration-Response Relationships between Ambient Ozone and Daily Mortality. *PLoS One* 2015; 10: e0129423.

2. Bero Bedada G, Raza A, Forsberg B, Lind T, Ljungman P, Pershagen G, Bellander T. Short-term Exposure to Ozone and Mortality in Subjects With and Without Previous Cardiovascular Disease. *Epidemiology* 2016; 27: 663-669.

3. Yin P, Chen R, Wang L, Meng X, Liu C, Niu Y, Lin Z, Liu Y, Liu J, Qi J, You J, Zhou M, Kan H. Ambient Ozone Pollution and Daily Mortality: A Nationwide Study in 272 Chinese Cities. *Environ Health Perspect* 2017; 125: 117006.

ISA, Executive Summary

Table ES-1: The order of outcomes in this table should reflect the order in the document: metabolic effects are discussed after cardiovascular effects.

Section IS.4.1 describes "Connections among health effects" in a potentially useful manner, and in a way that is not addressed in the individual appendices. However, the description on p. IS-20, of ozone effects in rats causing reductions in body temperature, BP, etc., as an example of "multisystem disruption", is somewhat confusing, because these responses do not occur in humans. The sentences following this refer to increased BP rather than decreased BP, which adds further confusion, especially since inceased BP is not an ozone response seen in the human clinical studies. This section needs to be re-thought and rewritten.

Appendix 3

Figure 3-1 provides an excellent synthesis of known and suspected biological pathways mediating ozone respiratory health effects. Some suggestions for further refinement:

1. Altered heart rhythm is included here, which is obviously not strictly a respiratory response. But other nonrespiratory links are not included here, that are consequences of ANS modulation and stress responses, including systemic inflammation and metabolic processes. This seems to be an inconsistency. Would remove altered heart rhythm from this figure for consistency.

2. Impaired host defense is shown linked solely with oxidative stress, but other pathways, for which there is evidence, are likely contributing, including airway injury, morphologic airway changes, and stress responses (elevated cortisol). Consider moving this box one column to the right, ungroup from morphologic changes and allergic responses and show as one of the downstream effects.

3. The pathway indicating that adrenal effects mediate airway injury/inflammation is based on a single study in rats (Miller et al 2016b). This finding runs counter to physiologic expectations (adrenal mediated stress response would be expected to follow acute inflammation/injury, not mediate it) and there is no evidence to support that this occurs in humans. Without further confirmation in additional studies or other species, or support of this directionality in humans, suggest making this line dotted.

P. 3-14, last paragraph, add Frampton et al. 2015 [1] to the list of new studies of lung function effects in the range of 100-300 ppb. This study included both GSTM1 sufficient and null subjects, and showed no effects of GSTM1 gene status on lung function responses.

P. 3-18, Cigarette Smoking. This section summarizes the Bates et al. 2014 study as showing similar lung function responses between smokers and nonsmokers, and indicates that this finding differs from previous studies. But the smokers in the Bates study were so-called "light" smokers, on average smoking about ½ pack per day for 6 years, for a total of 3 pack-years. This likely explains the difference from prior studies, which involved subjects with greater tobacco use, and this should be noted in the summary. For example, in Frampton et al. 1997 [2], one of the studies demonstrating significantly reduced lung function effects in smokers compared with never-smokers, only smokers of at least 1 ppd for a minimum of 3 years were included. The mean pack-years of smoking was 12.8. It should also be noted that, while ozone-induced lung function decrements are attenuated in smokers, lung inflammation is not [3], and oxidative stress may actually be increased [4]. This is an example of a situation where adverse respiratory effects of ozone may be occurring in the absence of lung function changes.

3.1.4.2.2 Animal Toxicological Studies, p. 3-23. Symptoms by definition are self-reported, and animals are obviously unable to report symptoms. It should be more clearly pointed out here that symptoms cannot be assessed in studies of rodents. Cough, or any other change in respiratory status, when reported by an observer, is a sign or an observation, not a symptom. It is only a symptom when reported by the individual experiencing it.

P. 3-38, line 1-2, Integrated summary. Change "FEV1" to FVC here. FEV1 is affected by changes in both volume (FVC, restrictive) and airways obstruction (FEV1/FVC).

P. 3-46, line 8. "These effects include sensory and pulmonary irritation…" The distinction here between "sensory" vs "pulmonary" irritation doesn't make sense. Pulmonary irritant responses have major sensory components. This phrase appears to be taken straight from the Hansen et al. 2016 abstract, but the terminology used in that abstract is not reflective of airway physiology. Sensory vagal-mediated inputs are important throughout the respiratory tract. The upper-lower airway distinction here is

incorrect, and is irrelevant to the point being made in this summary. The Hansen et al. study examines pulmonary outcomes, not upper airway responses.

Long term respiratory effects

The first paragraph of section 3.2.1, which includes a summary of the findings from the 2013 Ozone ISA, should include the limitations and uncertainties at that time that precluded a determination of "causal" for long-term respiratory effects.

Appendix 4

Figure 4-1 provides an excellent representation of the pathways, and evidence supporting them, leading to potential cardiovascular outcomes related to ozone exposure.

4.1.9.2, p. 4-23, 2nd bullet point. The description of the Arjomandi 2015 study, which is a clinical study, is written as if describing an epidemiology study. This paragraph should be re-written to indicate that subjects were exposed to clean air and 2 concentrations of ozone for 4 hours, with intermittent exercise, with HRV measured before and at intervals after exposure. In such a controlled and blinded experimental exposure, the changes can reasonably be described as effects of the exposure, rather than associations.

Table 4-4. The study by Frampton et al. 2015 did not assess LVDP. The cardiac function outcomes were cardiac index, stroke-volume index, and left ventricular ejection time. It is perhaps worth mentioning that these measures were obtained via impedance cardiography, rather than directly or via echocardiography.

Table 4-19. The study of Rich et al. 2018 measured SBP as well as DBP.

Table 4-26, Study-specific details from controlled human exposure studies of coagulation. This table should include Frampton et al. 2017 [5], which examined a number of coagulation parameters, without significant effects.

Table 4-29, systemic inflammation and oxidative stress markers. Add Frampton et al. 2017 [5] here as well.

Appendix 5 - Metabolic Effects

New determination of "likely to be causal". This determination is driven by the animal toxicology, which is largely limited to rodents. The animal data on glucose and insulin effects are robust. But the extrapolation of the findings to humans is in question. There appear to be no primate studies. The epidemiological evidence is sparse and inconsistent, without any evidence of adverse clinical outcomes related to metabolic effects. One human clinical study (Miller 2016a) showed no effects on insulin levels or HOMA-IR, but did find acute increases in stress hormones in response to ozone exposure. It is as yet

unconfirmed. While the animal studies provide plausibility, the sparse epi and human clinical data do not justify the "likely" determination. "Suggestive" appears to be more appropriate.

One question that should be considered is whether a brief stress response, in the absence of symptoms or other consequences, constitutes an adverse health effect. This could be considered a physiological response to a variety of stimuli. For example, it can occur in response to exercise.

Nevertheless, given the potential importance of these effects for human health, and in consideration of the current epidemics of obesity and diabetes, this represents an area of urgent research need.

5.1.4.1. Obesity animal toxicology studies: Some of the studies summarized here are relevant to obesity as a risk factor, in other words, whether obesity as a subject characteristic enhances ozone responses: pulmonary, CV, or other. Descriptions of these studies should be moved to the appropriate section on risk factors. The issue being considered in this section is whether ozone alters metabolic functions including body weight, BMI, body composition, caloric intake, glucose metabolism, lipid metabolism, stress responses, etc. The mixing of these two concepts is confusing and perhaps misleading. The sentence in this paragraph starting on line 12 describes what this section should be about: "Recent toxicological studies provided some evidence that ozone may impair metabolism and affect body weight, BMI, and body composition, as well as effect [sic] caloric intake."

The new evidence for metabolic effects does support the feasibility of ozone CV effects, given the strong link between the two.

P. 5-14: "Acute-phase liver proteins, such as CRP, can act as sensors of liver function". This is not accurate. CRP is made in the liver, and is a marker of systemic inflammation. Its production is driven by interleukin-6, released by a variety of cells during inflammation. Although produced in the liver, it is not considered a clinically useful marker of liver function.

5.1.5.3.4, Summary, p. 5-17. "Elevated circulating stress hormones are consistently observed in animal models and in controlled human exposure studies after short-term ozone exposure." This should be "in a single human controlled exposure study".

The last sentence of this summary statement ("Thus, neuroendocrine stress activation is essential to the development of adverse metabolic outcomes after short-term ozone exposure.") is overly broad and not completely supported by the described (adrenalectomy) studies.

5.1.5.4., p. 5-18. Serum lipids. The description of the Chen et al 2016a study is unclear, and it seems incorrect. According to the Abstract, this study deals with changes in lung function and nasal inflammation among schoolchildren. Was the reference intended to be Chen 2016b?

5.1.8., p. 5-23. Ketone bodies as a "marker" of diabetes is not accurate. Ketone bodies are also a "marker" of starvation or consuming a ketotic (low carb) diet. It is more accurately a marker of metabolic stress or perturbation with regards to glucose utilization. It does go up with diabetic ketoacidosis and can be considered a marker of that condition, but not of diabetes in general. Transient elevation of ketone bodies does not mean a person has or will get diabetes.

Table 5-1. “Consistent epidemiologic evidence” is inappropriate given there is only one study supporting it.

Appendix 6 - Mortality

Table 6-1, under “Key Evidence”, the statement is made, “Animal toxicological and controlled human exposure studies do not provide consistent evidence of potential biological pathways.” Actually, the experimental animal evidence for CV effects is fairly robust and convincing. It was mostly the inconsistency in the human studies and the relative lack of CV morbidity studies that led to the change in causality determination.

6.2.7, Summary and Causality Determination, Long-term total mortality, p. 6-40. The following statement in this section contrasts with previous text and the overall conclusions: “There is coherence across the scientific disciplines (i.e., animal toxicology, controlled human exposure studies, and epidemiology) and biological plausibility for ozone-related cardiovascular (Appendix 4) and respiratory (Appendix 3) endpoints, which lend some additional support to the ozone-mortality relationship.” The point is made repeatedly earlier in the ISA that the clinical studies are inconsistent with regard to CV effects. This sentence needs to be reconsidered and harmonized with the rest of the document.

Appendix 7 - Other health endpoints

Nervous system effects. Apparently included in this are the effects on the pulmonary irritant receptor/autonomic pathways that are well-established pulmonary effects in both animals and humans. Consideration should be given to separating this, and having this section include effects beyond the pulmonary irritant response loop, perhaps limiting it to brain, cognitive, and behavioral effects. Otherwise this category is causal based on the known local pulmonary neurological effects.

P. 7-42, line 21, “reproductive” effects should presumably be “nervous system” effects here.

References

1. Frampton M, Pietropaoli A, Dentler M, Chalupa D, Little E, Stewart J, et al. Cardiovascular effects of ozone in healthy subjects with and without deletion of glutathione-S-transferase M1. Inhal Toxicol. 2015;27:113-9.

2. Frampton MW, Morrow PE, Torres A, Cox C, Voter KZ, Utell MJ. Ozone responsiveness in smokers and nonsmokers. Am J Respir Crit Care Med. 1997;155:116-21.

3. Torres A, Utell MJ, Morrow PE, Voter KZ, Whitin JC, Cox C, et al. Airway inflammation in smokers and nonsmokers with varying responsiveness to ozone. Am J Respir Crit Care Med. 1997;156:728-36.

4. Voter KZ, Whitin JC, Torres A, Morrow PE, Cox C, Tsai Y, et al. Ozone exposure and the production of reactive oxygen species by bronchoalveolar cells in humans. Inhal Toxicol. 2001;13:465-83.

5. Frampton M, Balmes J, Bromberg P, Stark P, Arjomandi M, Hazucha M, et al. Multicenter Ozone Study in oldEr Subjects (MOSES). Part 1. Health Effects Institute Research Report 2017;No. 192.

Dr. Ronald J. Kendall

The "External Review Draft" integrated science assessment for ozone and related photochemical oxidants (hereafter referred to as "draft ISA") prepared by the United States Environmental Protection Agency's (U.S. EPA) National Center for Environmental Assessment – Research Triangle Park Division (NCEA – RTP) as part of EPA's ongoing review of the primary (health-based) and secondary (welfare-based) National Ambient Air Quality Standards (NAAQS) for Ozone was released on September 26, 2019. The EPA staff were directed by EPA Administrator Andrew Wheeler to "continue progress on the review of the NAAQS for ground level ozone through production of the Draft Ozone ISA and accelerating the development of a Draft Ozone Policy Assessment so that both documents could be delivered for CASAC and public review by October 2019. The present preliminary comments by Dr. Ronald J. Kendall will firstly address Appendix 8 of the draft ISA for ozone and secondly will address Appendix 9.

To supplement the standardized charge question and guide the scientific review of this ISA, the EPA has identified these additional areas for Clean Air Scientific Advisory Committee (CASAC) review and comment.

Ecological Effects of Ozone (Appendix 8)

> *Please comment on the identification, evaluation, and characterization of the available scientific evidence from studies of ecological effects of ozone, and the application of information from these studies, as presented in Appendix 8 to inform causality determinations for these welfare outcomes.*

First of all, determinations are made about causation by evaluating evidence across scientific disciplines and are based on judgements of consistency, coherence and biological plausibility of observed effects, as well as related uncertainties. It was noted that the ISA used a formal causal framework to classify the "weight of the evidence" using a five level hierarchy that characterized the evidence that forms the basis of causality determinations for welfare effect categories of a "causal relationship" or a "likely to be causal relationship" or describe instances where causality determination has changed (i.e., "likely to be causal" changed to "suggestive of, but not sufficient to infer a causal relationship"). Other relationships between ozone and welfare effects include "suggestive of, but not sufficient to infer" and "inadequate".

There are 12 causality determinations for ecological effects of ozone that are generally organized from the individual-organism scale to the ecosystem scale presented in Figure ES-5 in the ISA. To summarize the findings of the 2013 Ozone ISA, five are causal relationships (i.e., visible foliar injury, reduced vegetation growth, reduced crop yield, reduced productivity, and altered below ground biogeochemical cycles), and two are likely to be causal relationships (i.e., reduced carbon sequestration and altered ecosystem water cycling). One of the endpoints, alteration of terrestrial community composition, has now been concluded to be a "causal relationship" wherein the 2013 Ozone ISA this endpoint was classified as "likely to be causal". Three new endpoint categories (i.e., increased tree mortality, alteration of herbivore growth and reproduction, alteration of plant-insect signaling) not evaluated in the 2013 Ozone ISA are all determined to have a "likely to be causal relationship" with ozone. Plant

reproduction, previously considered as part of the evidence for growth effects is now a stand-alone causal relationship as illustrated in Figure ES-5.

Visible foliar injury from ozone exposure has been well characterized and documented over decades involving many trees, shrubs, herbaceous and crop species in using both long-term field studies and laboratory approaches. Even more recent experimental evidence continues to show consistent association between visible injury and ozone exposure supporting a "causal relationship" between ozone and visible foliar injury. Consistent with the 2013 Ozone ISA, there is a "causal relationship" between ozone and reduced plant growth and a "causal relationship" between ozone and reduced crop yield and quality. In the 2013 Ozone ISA, EPA considered reproduction in the same category with plant growth. Increased information of plant reproduction (such as flower number, fruit number, fruit weight, seed number, rate of seed germination) and evidence for direct negative effects on reproductive tissues, as well as for indirect negative effects (resulting from decreased photosynthesis and other whole plant physiological changes) warrants a special causality determination of a "causal relationship" between ozone exposure and reduced plant reproduction. Since the 2013 Ozone ISA, large-scale statistical analysis of many factors concluded that county-level ozone concentrations averaged over the study period significantly increased tree mortality and many plant functional types. This evidence, combined with observations of long-term declines of conifer forests in several high ozone regions and new experimental evidence that sensitive genotypes of, particularly, aspen trees have increased mortality with ozone exposure, support a "likely to be causal relationship" between ozone exposure and tree mortality.

In addition to the direct effects of ozone on plants, ozone can alter ecological interactions between plants and other species, including herbivores that may consume ozone-exposed vegetation. Some recent evidence of insect herbivores in previous ozone assessments and new studies covering a range of species provide collective evidence that supports a "likely to be causal relationship" between ozone exposure and altered herbivore growth and reproduction. Many plant-insect interactions are mediated between volatile plant signaling compounds, which plants use to signal other members within an ecological community. New evidence from multiple studies show altered/degraded emissions of chemical signals from plants and reduced detection of plant signaling compounds by insects. Therefore, the collective evidence supports "a likely to be causal relationship" between ozone exposure and alteration of plant-insect signaling.

At the ecosystem scale, ozone caused suppression of plants' photosynthesis which can lead to reduced ecosystem carbon content. Consistent with the conclusions of the 2013 Ozone ISA, there is a "causal relationship" between ozone exposure and reduced productivity and a "likely to be causal" relationship between ozone and reduced carbon sequestration. Recent evidence continues to support a "causal relationship" between ozone exposure and the alteration of below ground biogeochemical cycles. We know ozone can affect water use in plants through several mechanisms and ultimately affect plant evapotranspiration, which may in turn lead to possible effects on hydrogeological cycling. Evidence continues to support the conclusion of the 2013 Ozone ISA that there is a "likely to be causal relationship" between ozone and alteration of ecosystem water cycling. Alteration of community composition of some ecosystems, including conifer forests, broadleaf forests and grasslands, and altered fungal and bacterial communities in the soil reported in the 2013 Ozone ISA is augmented by additional evidence for effects in forests and grassland communities indicating a change in the causality

determination to a "causal relationship" between ozone exposure and altered terrestrial community composition of some ecosystems.

The summary of causality determinations for ecological effect are summarized as follows:

1. Conclusions from the 2013 Ozone ISA that support the seven conclusions of causality in the current 2019 Ozone ISA include 1) visible foliar injury, 2) reduced vegetation growth, 3) reduced plant reproduction, 4) reduced yield and quality of agricultural crops, 5) reduced productivity in terrestrial ecosystems, 6) alteration of below ground biogeochemical cycles, and 7) alteration of terrestrial community composition.

The "weight of the evidence" appears to strongly support the previous conclusions from the 2013 Ozone ISA subsequently identified in the conclusions in the current ISA. The summary of five causality determinations for ecological effects in the 2019 Ozone ISA, which build on the conclusions from the 2013 Ozone ISA, include the following:

1. Reduced plant reproduction from no "separate causality" to a "causal relationship" with ozone exposure,
2. Increased tree mortality "causality not assessed" and changed to "likely to be a causal relationship",
3. Alternation of herbivore growth and reproduction changed from "causality not assessed" to "likely to be causal relationship",
4. Alteration of plant-insect signaling "causality not assessed" changed to "likely to be a causal relationship",
5. Alteration of terrestrial community composition changed from "likely to be a causal relationship" to "causal relationship".

For these five causality determinations for ecological effects that have changed in terms of conclusions in the current ISA from the conclusions from the 2013 Ozone ISA will be more fully evaluated in terms of preliminary comments from the initial review of these data.

Appendix 8 "Ecological Effects" in the 2019 Ozone ISA evaluates the relevant scientific information on ecological effects as part of the review of the air quality criteria for ozone and other photochemical oxidants and to help form the scientific foundation for the review of the secondary NAAQS for ozone. This Appendix serves as an update to Chapter 9 of the 2013 Ozone ISA (U.S. EPA, 2013). The majority of the evidence for ecological effects has been for vegetation. Effects at the individual plant level can result in broad ecosystem-level changes, such as productivity, carbon storage, water cycling, nutrient cycling, and community composition. The current ISA has adopted the use of the Population, Exposure, Comparison, Outcome, and Study design (PECOS) tool to further define the scope of the current review by conveying the criteria for inclusion or exclusion of studies. The units of study as defined in the PECOS for ecological effects of ozone are the individual organism, species, population, community, or ecosystem. It should be noted that all studies included in the 2019 Ozone ISA were conducted at concentrations occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations observed in the U.S. For ecological endpoints for which the 2013 Ozone ISA concluded that the evidence was sufficient to infer a causal relationship (i.e., foliar injury,

vegetation growth, ecosystem productivity, yield and quality of agricultural crops, below ground biogeochemical cycling). These were fully evaluated in the 2019 Ozone ISA. In terms of new determination or change in causality from the 2013 Ozone ISA, the following causality determinations for ecological effects of ozone will be addressed in the current review. At the community level, biodiversity in terms of terrestrial community composition is now "causal", and species interactions including plant-insect signaling is a new determination and "likely to be causal". In addition, tree survival is changed to "likely causal" and growth of insect herbivores feeding on ozone-affected plants is "likely causal". The plant reproduction endpoint is now separate from plant growth and a new determination as "causal" and new determination of growth and reproduction is "likely to be causal" is assigned to insect herbivores affected by ozone. All causality determinations or changes in causality determination from the 2013 Ozone ISA will be thoroughly considered in the present series of comments. The current review only evaluates studies conducted in North America. In the PECOS for ecological effects, relevant study designs include laboratory, greenhouse, field, gradient, open top chamber (OTC), free air carbon dioxide enrichment (FACE), and modeling studies.

Visible Foliar Injury in Biomonitoring

In the 2013 Ozone ISA, the evidence was sufficient to conclude that there is a causal relationship between ambient ozone exposure and the occurrence of ozone-induced visible foliar injury on sensitive plant species across the U.S. (U.S. EPA, 2013). Visible foliar injury from exposure to ozone has been well characterized and documented on many tree, shrub, herbaceous, and crop species through research beginning in 1958. Ozone-induced visible foliar injury is considered diagnostic because it has been experimentally induced and it is considered a bioindicator for ozone exposure in plants. As described in the PECOS tool, the scope for new evidence reviewed in the section limits studies to those conducted in North America at concentrations occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations. Experimental evidence continues to show a consistent association between visible injury and ozone exposure in plants. Since the 2013 Ozone ISA, several studies have further characterized modifying factors 1) additional field studies have shown dry periods tend to decrease the incidence and severity of ozone-induced visible foliar injury, 2) data used in additional species from greenhouse studies add to the evidence that sensitivity to ozone varies by the time of day in plants, 3)phenotypic variation of foliar sensitivity to ozone has been observed, 4) in OTC exposure (mean 12 hour ozone concentration of 37 ppb for 118 days) foliar injury to loblolly pine seedlings were not related to seedling inoculation with root-infecting fungi (Chieppa et al, 2015).

Since the 2013 Ozone ISA, several additional studies have been conducted on bioindicator species:

1. Cutleaf coneflower is an ozone bioindicator species native to Great Smokey Mountains National Park,
2. Tree of heaven, an established invasive species found widely across the U.S., has been identified as an effective ozone bioindicator species by the National Park Service and Forest Service (Smith et al, 2008; Kohut, 2007).

In greenhouse exposures, foliar injury occurred at 8 hour average ozone exposure levels of 60 – 120 ppb with greater injury corresponding to higher exposure (Seiler et al, 2014). As noted in the 2013 ISA, visible foliar injury usually occurs when sensitive plants are exposed to elevated ozone concentrations in a predisposing environment. A major modifying factor for ozone-induced visible foliar injury is the

amount of soil moisture available to a plant during the year that the visible foliar injury is being assessed. This is because the lack of soil moisture generally decreases stomatal conductance of plants and, therefore, limits the amount of ozone entering the leaf that can cause injury. Visible foliar injury from ozone exposure has been well characterized for decades using both long-term field studies and laboratory approaches. Since the 2013 Ozone ISA, new research on bioindicator species and the further characterization of modifying factors have provided further support for the effects. New information is consistent with the conclusions of the 2013 Ozone ISA that the body of evidence is sufficient to infer a causal relationship between ozone exposure and visible foliar injury. With the decades of research, both in field observation as well as experimental studies related to the foliar injury endpoint, the body of evidence remains very strong to infer a causal relationship between ozone exposure and visible foliar injury.

Plant Growth

In the 2013 Ozone ISA, the evidence was sufficient to conclude that there is a causal relationship between ambient ozone exposure and reduced growth of native woody and herbaceous vegetation (U.S. EPA, 2013). In the 2013 Ozone ISA, it was concluded there is strong and consistent evidence that exposure to ozone decreases photosynthesis and growth in numerous plant species. The evidence available at that time and now discussed in the 2019 Ozone ISA found that ambient ozone concentrations caused decreased growth (measured as biomass accumulation) in annual, perennial, and woody plants inclusive of crops, annuals, grasses, shrubs, and trees. A meta-analysis by Wittig, et al (2009) found that the average ozone exposures of 40 ppb significantly decreased annual total biomass by 7% across 263 studies. Biomass declines were linked to reductions in photosynthesis (U.S. EPA, 2013), which are consistent with cumulative plant uptake of ozone into the leaf (Wittig et al, 2007). Further, there is evidence ozone may change plant growth patterns by significantly reducing carbon allocated to roots in some species. Since the 2013 Ozone ISA, there is more evidence from experimental studies that support detrimental effects of ozone on plant growth:

1. Results from aspen-only stand at the Aspen FACE experiment in Wisconsin showed a decrease of 12 – 19% in the relative growth rate of 3 of 5 genotypes of aspen studied,
2. When site level results from Aspen FACE experiment were scaled up using the forest landscape model (LANDIS II), ozone was found to significantly reduce landscape biomass,
3. In meta-analysis of 9 studies examining intra-specific variation in juvenile tree growth under elevated ozone, found that elevated ozone generally reduced photosynthetic rate as well as height growth and stem volume,
4. A study using the invasive Chinese tallow tree suggests ozone response may be genotype-specific,
5. Using model simulation coupled with established U.S. EPA ozone exposure response functions in seedlings, estimated relative biomass loss at 2.5% for Ponderosa pine and 2.9% for aspen, and
6. In another estimation of biomass loss of adult trees across the U.S. for modeled ozone values, eastern cottonwood and black cherry showed high sensitivity.

In addition to these studies, there is a recent global scale synthesis of published ozone exposures studies that document reductions in biomass due to ozone exposure in over 100 plant species (Bergmann et al, 2017). In the 2019 Ozone ISA, there is strong scientific evidence sufficient to conclude that there is a

causal relationship between ambient ozone exposure and reduced growth of native woody and herbaceous vegetation.

Reduced Plant Reproduction

In the 2013 Ozone ISA, reduced plant reproduction was not separated for causality determination and was included with plant growth. However, in the 2019 Ozone ISA, reduced plant reproduction is scientifically defended for a causal relationship between plant reproduction metrics and exposure to ozone. In fact, the recent literature shows that across most plant reproduction metrics (such as flower number, fruit number, fruit weight, seed number, and rate of seed germination) with elevated exposure concentrations that ozone has significant negative effects on plant reproduction. In a first of its kind study, Leisner and Ainsworth (2012) conducted a quantitative meta-analysis to assess the general magnitude and direction of the effects of ozone exposure on plant reproduction. In experiments that used ambient air as the control, average fruit weight decreased 51% (at an average exposure of 98 ppb), which was the largest effect observed in this part of the meta-analysis, and seed number decreased approximately 10% (at an average exposure of 68 ppb). In studies with ozone-sensitive species of clover, Sanz et al (2016) showed that reproduction was reduced significantly with increasing ozone exposure. Gillespie et al (2015) isolated the effects of ozone on particular reproductive tissues of tomato. Pollen grains exposed to ozone have significantly reduced germination and pollen tube growth in vitro. Reductions in pollen viability is, for example, and extremely important plant reproduction metric.

Timing of ozone exposure relative to reproductive development stages can affect reproductive outcomes in some cases. Flowers exposed to ozone early in their development tended to produce shorter fruits than flowers exposed later in their development. There appears to be adequate information, particularly from the quantitative meta-analysis reported by Leisner and Ainsworth (2012) supporting a causal relationship between ozone exposure and reduced plant reproduction. The strength of the scientific support for supporting a "causal relationship" is not as strong as with visible foliar injury and reduced vegetation growth as previously reviewed in the current comments. However, with the separate category of reduced plant reproduction, it can be concurred that causality does exist between ambient ozone exposure and this plant metric. It has been shown that diverse metrics of plant reproduction decline under ozone concentrations occurring in either the environment or under experimental conditions within an order of magnitude of recent concentrations. Metrics of plant reproduction, fruit number and fruit weight, show reductions under increased ozone when combined across species for ozone concentrations that span 40 to >100 ppb. Finally, experimental ozone exposure at multiple experimental settings (such as in vitro, whole plants in the laboratory, whole plants and/or reproductive structures in the greenhouse, and whole plant communities in field settings) convincingly show ozone independently reduces plant reproduction. I concur that previous evidence and new evidence reviewed here is sufficient to infer a "causal relationship" between ozone exposure and reduced plant reproduction.

Plant Mortality

In the 2013 Ozone ISA, causality was not assessed for increased tree mortality involving ozone exposure. The conclusions in the 2019 Ozone ISA is that there is "likely to be a causal relationship" between ozone exposure and plant mortality. Several new studies examine the impacts of ozone exposure on plant mortality that included the fraction of individuals in a population that die over a given timeframe. These experiments were focused on tree species demonstrating ozone exposure can affect

tree mortality. For instance, in the Aspen FACE experiment, the survival of sensitive aspen genotypes 271 and 259 declined significantly between 1997 and 2008 under elevated ozone exposures (Moran and Kubiske, 2013). In addition, Dietz and Moorcroft, 2011) conducted a large-scale analysis of factors contributing to annual mortality of trees and functional types in the forests of the eastern and central U.S. In their analysis, ozone was ranked 9th on a list of 13 factors that forests were sensitive to an ozone's effects with similarly magnitude to that of precipitation. Mortality in 8 out of 10 plant functional types were significantly correlated with ozone 8 hour max exposures. Therefore, studies of tree mortality indicate that ozone affects this endpoint. Studies linking ozone and tree mortality are consistent with known and well-established individual plant level mechanisms that explain ozone phytotoxicity, including variation and sensitivity and tolerance based on age class, genotype, and species. Experimentally elevated ozone exposures has been shown to increase mortality in sensitive Aspen genotypes. Considering the previous evidence and new evidence reviewed in the 2019 Ozone ISA, it is sufficient to infer a "likely to be causal relationship" between ozone exposure and tree mortality.

Reduced Crop Yield and Quality

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a "causal relationship" between ozone exposure and reduced yield and quality of agricultural crops (U.S. EPA, 2013). The detrimental effect of ozone on crop production has been recognized since the 1960s and there is a large body of research that has subsequently characterized decreases in yield and quality of agricultural crops. As described in the PECOS tool, the scope of new evidence reviewed in this section are limited to studies conducted in North America at ozone concentrations occurring in the environment or experimental ozone concentrations within an order of magnitude of research concentrations.

For soybeans, additional studies in Illinois report decreased seed/crop yield (Leisner et al, 2017). A linear decrease in soybean yield was observed across two growing seasons at the rate of 37-39 kg/ha per ppb cumulative ozone exposure over 40 ppb. For wheat, meta-analysis using data from the U.S. and other countries provide further supporting evidence that current levels of ambient ozone decrease growth, quality, and yield (Pleijel et al, 2018). New studies in non-soybean legumes include evaluation of biomass and seed yield in ozone-exposed snap bean under high- and low-vapor pressure deficit conditions (Fiscus et al, 2012). U.S. modeling studies in the 2013 Ozone ISA found that ozone generally reduced crop yield and that different crops showed different sensitivity to ozone (Avnery et al, 2011). Newly available regional and national scale analyses of ozone effects on major crops in the U.S., including soybean, wheat, and maize have further enabled characterization and quantification of yield losses (McGrath et al, 2015).

The relationship between ozone exposure and reduced crop yield is well established in the scientific literature and continues to be an active area of research with many new scientific papers being published since the 2013 Ozone ISA. Recent advances in characterizing ozone effects on U.S. crop yield include further geographic and temporal refinement of ozone sensitivity in national scale estimates of maize and soybean losses from ozone based on actual yield data. The new scientific information published is consistent with the conclusions of the 2013 Ozone ISA that the body of evidence is sufficient to infer a "causal relationship" between ozone exposure and reduced yield and quality of agricultural crops.

Herbivores: Growth, Reproduction, and Survival

In the 2013 Ozone ISA, there was no causality determination between ozone exposure and effects on herbivores. We know that ozone exposure can lead to changes in plant physiology, such as by modifying the chemistry and nutrient content of leaves (U.S. EPA, 2013). These changes can have significant effects on herbivore physiology and behavior. There was no consensus in the 2013 Ozone ISA on how insects and other wildlife respond to elevated ozone. Since that review, additional research has been published for more herbivorous insects as well as a few mammalian herbivores at various levels of ozone exposure. As described in the PECOS tool, the scope of this review includes studies in which alterations in invertebrates and vertebrate responses were measured in individual species or at the population and community levels as related to concentrations of ozone occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations. In the 2013 Ozone ISA, a meta-analysis that included 16 studies published on insect herbivore species between 1996 and 2005 found that elevated ozone decreased development time and increased pupil mass in insect herbivores with more pronounced effects occurring with longer durations of ozone exposure (Valkama et al, 2007). Since the 2013 Ozone ISA, there is new evidence for endpoints related to growth, reproduction, and survival in insect herbivores encompassing the orders Coleoptera, Hemiptera, and Lepidoptera. With the available science reported in the 2019 Ozone ISA regarding the effects of ozone on growth, reproduction, and survival of, particularly, insect herbivores substantial new information has been made available in order to assess a causality relationship. In addition, population and community level responses reveal that changes in host –plant quality resulting from elevated ozone can alter the population density and structure of associated insect herbivore communities ultimately affecting ecosystem processes (Cornelissen, 2011). Recent studies reviewed in the 2019 Ozone ISA include multiple experimental studies conducted by many research groups that expand the evidence base for the effects of elevated ozone on growth and reproduction in herbivores. It is recognized that while effects were observed there remains a more limited number of studies on the effects of ozone on survival and population/community level responses. Recognizing that since the 2013 Ozone ISA and with increased research efforts on herbivore response to plants impacted by ozone, a new causality determination appears justified that the body of evidence is sufficient to infer a "likely to be causal" relationship between ozone exposure and alteration of herbivore growth and reproduction.

Alteration of Plant-Insect Signaling

In the 2013 Ozone ISA, there was "no causality" determination between ozone exposure and alteration of plant-insect signaling. Plants signal to other ecological community members through the emission of volatile plant signaling compounds (Blande et al, 2014). Each signal emitted by plants has an atmospheric lifetime and a unique signature comprised of different ratios of individual hydrocarbons that are susceptible to atmospheric oxidants, like ozone (Yuan et al, 2009). Insects and other fauna discriminate between chemical signals of different plants. As described in the PECOS tool, the scope in the 2019 Ozone ISA for considering plant-insect signaling include studies that assess altered plant-insect signaling in response to concentrations of ozone occurring in the environment or experimental ozone concentrations within an order of magnitude of recent concentrations. Under conditions of elevated ozone the degradation of plant signaling compounds resulted in bumble bees orienting significantly less towards floral scent queues and exhibiting preference for artificial flowers closer to the ozone source (Farre-Armengol et al, 2015). As reported previously, herbivorous insects use plant signaling compounds to locate suitable host plants and ozone can alter these interactions (Blande et al, 2010). In

chamber studies, elevated ozone reduced the ability of insect herbivores to find their plant host (Li et al, 2016). Striped cucumber beetles could not distinguish between clean air and air containing floral volatiles when the ozone concentration exceeded 80 ppb (Fuentes et al, 2013). In addition, plant defense responses include emission of plant-signaling compounds to attract predators and parasitoids that target the herbivores feeding on the plant. In studies reviewed in the 2013 Ozone ISA and new studies on parasitoid-host attraction show either reduced, enhanced, or unaffected behavior by elevated ozone (Cui et al, 2016). Altered plants signaling to natural enemies of herbivores disrupts predator-prey trophic interactions. The interaction of ozone (>50 ppb) with plant signaling compounds disrupts the production, emission, dispersion, and lifespan of these compounds. Considering the available evidence reported in the 2013 Ozone ISA and more recent research efforts while as well recognizing uncertainties around how chemical signaling responses observed in the laboratory translate to natural environments, the 2019 Ozone ISA makes a new causality determination that the body of evidence is sufficient to infer a "likely causal relationship" between ozone exposure and alteration of plant-insect signaling.

Reduced Productivity in Terrestrial Ecosystems

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a causal relationship between ozone exposure and reduced plant productivity. The terrestrial carbon cycle integrates processes at various scales, ranging to organelles to individuals to biomes (Chapin et al, 2002). Gross primary productivity, which is the influx of CO_2 from the atmosphere via photosynthesis at the ecosystem scale is fundamental to global carbon cycling. Since the 2013 Ozone ISA, two new studies have reported on the effects of ozone on gross primary productivity. Fares et al (2013) conducted statistical analysis of data to quantify the effect of ozone on carbon assimilation. In California, ozone decreased carbon assimilation by 12% in pine forests in the Sierra Nevada and by 19% in an orange grove in the Central Valley. Yue and Unger (2014) adopted the same ozone-damaged thresholds in their analysis that were used in previous models to assess ozone damage. What was learned was decreases in gross primary productivity as a result of ozone range from 1-14% and were greatest at sites showing both high stomatal conductance and high growing season ozone concentrations. Carbon assimilated into plant tissue via photosynthesis is either respired or contributes to net primary productivity, which is often measured as the rate of plant biomass accumulation. While much of the research published since 2013 Ozone ISA is confirmatory, other work has provided new mechanistic insight into the effects of ozone on net primary productivity. Evidence of the effect of ozone exposure in ecosystem productivity comes from many different experiments with different study designs in a variety of ecosystems and models. New information is consistent with conclusions of the 2013 Ozone ISA that the body of evidence is sufficient and increasing to infer a "causal relationship" between ozone exposure and reduced ecosystem productivity.

Reduced Carbon Sequestration in Terrestrial Ecosystems

Terrestrial carbon sequestration is the sum of carbon contained within biomass and soil within a defined ecosystem typically quantified on a multi-year scale (Koerner 2006). As in the 2013 Ozone ISA, most assessments of the effects of ozone on terrestrial carbon sequestration are from model simulations. However, an assessment of the effect of ozone on ecosystem carbon content at the Aspen FACE experiment was published in 2014. At the conclusion of the Aspen FACE experiment after 11 years of fumigation, Talhelm et al (2014) observed that elevated ozone decreased ecosystem carbon content (in plant biomass, litter, and soil carbon to 1 m in depth) by 9%. Total tree biomass carbon was 15% lower

under elevated ozone with decreased woody biomass counting for nearly all of the effect of tree biomass. The results from the Aspen FACE experiment and the model simulations provide further evidence that ozone can decrease ecosystem carbon sequestration. Although the decrease in net primary productivity were temporary in the Aspen FACE experiment, the 10% decrease in cumulative net primary productivity at Aspen FACE was associated with a 9% decrease in ecosystem carbon storage (Talhelm et al, 2014). The relationship between ozone exposure and terrestrial carbon sequestration is difficult to measure at the landscape scale. Most of the evidence regarding this relationship is from model simulations, although this endpoint was examined in a long-term manipulative chamber-less ecosystem experiment known as Aspen FACE, already described. Even with limitations, the result from the Aspen FACE experiment and supported by model simulation provide further evidence that is consistent with the conclusions of the 2013 Ozone ISA that the body of evidence is sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and reduced carbon sequestration in ecosystems.

Soil Biogeochemistry

The 2013 Ozone ISA concluded there is a "causal relationship" between ozone exposure and the alteration of below ground biogeochemical cycles (U.S. EPA, 2013). This causality determination was based on the body of evidence known at that time. The 2013 Ozone ISA (U.S. EPA, 2013) presented evidence that ozone alters multiple below ground endpoints, including root growth, soil food web structure, soil decomposer activities, soil respiration, soil carbon turnover, soil water cycling, and soil nutrient cycling. The new evidence since the 2013 Ozone ISA included in the 2019 Ozone ISA confirms ozone affects soil decomposition, soil carbon, and soil nitrogen. Soil carbon is often a mix of inorganic and organic forms of carbon, the latter may be from living and/or dead plant animal, fungal, and bacterial organisms. The effects of ozone on several aspects of soil carbon have been investigated. Ozone can alter the cycling of nitrogen in the soil via its direct effect on plants. Nitrogen is an important element to plant life as it is often the limiting nutrient from most temperate ecosystems. The 2013 Ozone ISA (U.S. EPA, 2013) documented mixed results of ozone effects on soil nitrogen pools and processes with results indicating no effect in meadow nitrogen biomass or potential nitrification and denitrification (Kanerva et al, 2006). While ozone was shown to increase nitrogen released from litter in a forest (Stoelken et al, 2010), ozone decreased gross nitrogen mineralization (Holmes et al, 2006) at Aspen FACE and nitrogen release from soil litter. The 2013 Ozone ISA presented evidence that ozone was found to alter multiple below ground endpoints, including root growth, soil food web structure, soil decomposer activities, soil respiration, soil carbon turnover, soil water cycling, and soil nutrient cycling. New evidence since the 2013 Ozone ISA (U.S. EPA, 2013) included in this assessment confirms ozone effects on soil decomposition, soil carbon, and soil nitrogen. Overall, the evidence does not change the conclusions from the 2013 Ozone ISA (U.S. EPA, 2013) and, therefore, suggests that ozone can alter soil biogeochemical cycling of carbon and nitrogen, although the direction and magnitude of these changes often depends on the species, site, and time of exposure. Currently, it is recognized that it does not appear to be a consistent exposure-response relationship. The body of evidence is sufficient to conclude that there is a "causal relationship" between ozone exposure and the alteration of below ground biogeochemical cycles.

Alteration of Terrestrial Community Composition

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and alteration of terrestrial community composition of some ecosystems (U.S. EPA, 2013). Ozone altered above ground plant communities, such as conifer forests, broadleaf forests, and grasslands and altered fungal and bacterial communities in the soil in both natural and agricultural systems (U.S. EPA, 2013). Ozone effects on individual plants can alter the larger plant community as well as the below ground community of microbes and invertebrates, which depend on plants as carbon sources. In the 2013 Ozone ISA, evidence of ozone effects on forest composition was drawn from the observational studies of conifer decline correlated with ozone exposure (Allen et al, 2007). New evidence suggests that ozone alters tree competitive interactions for nutrients, such as consistent with previous research on altered tree community composition at Aspen FACE showed that elevated ozone altered the relative competition for nutrients among aspen genotypes (Zak et al, 2012). Since the 2013 Ozone ISA, new studies extend the scope of evidence regarding forest community composition to include synthesis and models. In the 2013 Ozone ISA, there was evidence of ozone effects on grassland community composition in controlled experimental exposure studies, in models, and in reviews. Key new studies include experimental ozone exposures that allow evaluation of ozone effects on grassland community composition and analyses that explicitly include environmental or annual heterogeneity.

Even with microbes, the 2013 Ozone ISA documented effects of ozone on soil microbial communities with changes in proportions of bacteria or fungi as a result of experimental ozone exposures in grassland mesocosms, peatland mesocosms, and forest mesocosms. In addition, changes in soil microbial communities in agricultural systems was reported (Chen et al, 2010). Even with bacteria, the 2013 Ozone ISA found decreases in bacterial abundance in response to elevated ozone in meadows and forests mesocosms. There have been many new studies reported to assess the effect of elevated ozone on soil bacteria. The 2013 Ozone ISA found effects of ozone exposure on soil fungi (U.S. EPA, 2013). Studies found that ozone exposure decreased fungal biomass in meadow mesocosms, marginally increased fungal abundance in peatland mesocosms and altered fungal community composition in forest soils. Many new studies have evaluated the effects of ozone on fungi since the 2013 Ozone ISA. The 2013 Ozone ISA found evidence sufficient to conclude that there is a "likely to be causal relationship" between ozone exposure and the alteration of community composition of some ecosystems. Evidence of this relationship was presented for forest communities of trees, grassland communities of grasses, herbs, and legumes and soil microbial communities of bacteria and fungi. Recently published papers extend the evidence of each of these topics in the 2013 Ozone ISA.

In forests, previous evidence included correlation on studies across ambient gradients of ozone exposure that found effects of ozone on conifer trees, studies with controlled experimental exposure of trees that found effects of ozone on deciduous trees. Key new studies show that observational and experimental observations of ozone effects on tree species extend to affect regional forest composition in the Eastern U.S. (Wang et al, 2016). In grasslands, previous evidence included multiple studies from multiple research groups to show that elevated ozone shifts the balance among grasses, forests, and legumes. There are new studies that show ozone affected the ratio of grass to legume biomass (Gilliland e al, 2016). In soil microbial communities, previous evidence includes studies that found effects on the ratio of bacteria to fungi in soil communities as well as effects on community composition of mycorrhizal fungi. New studies confirm that elevated ozone alters soil microbial taxa, although as with previous

evidence, the strength and directional effects are not consistent across all ecosystems. The 2013 Ozone ISA presented multiple lines of evidence that elevated ozone alters terrestrial community composition, and recent evidence strengthens our understanding of the effects of ozone on plant communities while confirming that the effects of ozone on soil microbial communities are diverse. The body of evidence is sufficient to conclude that there is a "causal relationship" between ozone exposure and the alteration of community composition of some ecosystems.

Alteration of Ecosystem Water Cycling

In the 2013 Ozone ISA, the evidence was sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and the alteration of ecosystem water cycling (U.S. EPA, 2013). Plants are responsible for part of the ecosystem water cycling through root uptake of soil moisture and groundwater as well as transpiration through leaf stomata to the atmosphere. Changes to this part of the water cycle may in turn affect the amount of water moving through the soil, running off over land or through groundwater and flowing through streams. Ozone can affect water use in plants and ecosystems through several mechanisms, including damage to stomatal functioning and loss of leaf area, which may affect plant and stand transpiration. During the review of the 2013 Ozone ISA, there was debate on the assumption that ozone exposure consistently reduced stomatal conductance in plants. Several studies have found increased conductance, suggesting stomatal dysfunction in response to ozone exposure. However, other studies found ozone caused a loss of stomatal control, incomplete stomatal closure at night, and a decoupling of photosynthesis in stomatal conductance. There is mounting biologically relevant and statistically significant data from multiple studies showing the mechanisms of ozone effects on plant-water use in ecosystem water cycling (reduced leaf area, reduced leaf longevity, changes in root and branch biomass and architecture, changes in vessel anatomy, stomatal dysfunction, reduced sap flow). The most compelling evidence showing effects at the ecosystem level is from studies in Eastern U.S. forests and from the Aspen FACE. All of this new information supports the 2013 Ozone ISA and supports the conclusion in the 2019 Ozone ISA that the body of evidence is sufficient to conclude there is a "likely to be causal relationship" between ozone exposure and the alteration of ecosystem water cycling.

General Comments

1. I compliment the United States Environmental Protection Agency for the thoroughness and completeness of Appendix 8 entitled, "Ecological Effects" as part of the 2019 Ozone Integrated Science Assessment.
2. I agree with the "causality" determinations for the components for ecological effects considered in the 2019 Ozone Integrated Science Assessment including 1) visible foliar injury, 2) reduced vegetation growth, 3) reduced plant reproduction, 4) reduced yield and quality of agricultural crops, 5) reduced productivity in terrestrial ecosystems, 6) alteration of below ground biogeochemical cycles, and 7) alteration of terrestrial community composition. I agree with the "likely to be causal" determinations including 1) increased tree mortality, 2) alteration of herbivore growth and reduction, 3) alteration of plant-insect signaling, and 4) reduced carbon sequestration in terrestrial ecosystems, and 5) alteration of ecosystem water cycling
3. It is my impression that a thorough review and reporting of the scientific literature that has been generated since the 2013 Ozone ISA has been incorporated into the 2019 Ozone ISA.

4. In terms of the summary of causality determinations for ecological effects, I support the determinations made by the U.S. EPA as a function of the science availability and its interpretation.
5. Although historically the predominant ecological effects assessed with ozone exposure has been with vegetation, the current Appendix 8 "Ecological Effects", has at least some mention of terrestrial vertebrates, including rabbits and goats, and how they may respond to altered vegetation as a function of ozone exposure. I think that this area should be expanded in terms of alteration of individual plants and plant communities can disrupt terrestrial vertebrates, and not just invertebrates. Therefore, I recommend consideration of an expanded research plan to look at the implications of altered vegetation communities from ozone exposure and response to terrestrial vertebrate herbivores.
6. Although there is in depth consideration in other sections of the 2019 Ozone Integrated Science Assessment involving human health implications from ozone exposure, which are real and well-defined cause and effect relationships that have been scientifically studied a considerable length of time, nothing is mentioned with wildlife. In Appendix 8 "Ecological Effects", there is no mention whatsoever of wildlife toxicology implications for ozone exposure, although human health implications have been considerably considered in other parts of the 2019 Ozone Integrated Science Assessment. I recommend to at least consider and develop a research plan for a bird model that could be assessed in terms of the wildlife toxicology of ozone exposure in warm-blooded vertebrates. This would be essentially a "canary in the coal mine" concept for detecting toxic gasses by miners through a bird model. I think this same concept could be implemented utilizing an avian model for the study of ozone exposure in terrestrial warm-blooded non-human vertebrates (Kendall at el, 2010).

Climate Change (Appendix 9)

Please comment on the identification, evaluation and characterization of the available scientific evidence from studies of ozone effects on climate, and the application of information from these studies as presented in Appendix 9 to inform causality determinations for these welfare outcomes.

For effects on climate, changes in the abundance of tropospheric ozone disturbs the radiative balance of the atmosphere by interacting with incoming solar radiation and outgoing longwave radiation. This effect is quantified by radiative forcing, which is the perturbation in net radiation flux at the tropopause caused by a change in radiatively active forcing agent after stratospheric temperatures have readjusted to radiative equilibrium. Through this effect on the earth's radiation balance, tropospheric ozone plays a significant role in the climate system and increases in tropospheric ozone abundance contribute to climate change as addressed in the 2013 Ozone ISA. Recent evidence continues to support a causal relationship between tropospheric ozone and radiative forcing and a "likely to be causal relationship" via radiative forcing between tropospheric ozone and temperature, precipitation and related climate variables referred to as climate change in the 2013 Ozone ISA. New evidence comes from the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) (Myhre et al, 2013) and supporting references. As thoroughly discussed in the 2019 Ozone ISA, none of the new studies indicate a change to either causality determination included in the 2013 Ozone ISA. In terms of effects of tropospheric ozone and climate change, radiative forcing remains a "causal" relationship and temperature, precipitation, and related variables maintain a "likely to be causal" relationship. Consistent

with previous estimates in the 2013 Ozone ISA, the 2019 Ozone ISA is consistent with previous estimates, the effect of tropospheric ozone on global surface temperature through its impact on radiative forcing continues to be estimated at roughly 0.1 to 0.3°C since industrial times. While the warming effect of tropospheric ozone in the climate system is established, precisely quantifying changes in surface temperature due to tropospheric ozone changes along with related climate effects requires complex climate simulations. There are current limitations in climate modeling tools that need to be recognized and the need for more comprehensive observational data on these effects represent sources of uncertainty in quantifying the precise magnitude of climate response to ozone changes (Myhre et al, 2013). All of this evidence reinforces the "likely to be causal" relationship between tropospheric ozone and temperature, precipitation, and related climate variables which was referred to as "climate change" in the 2013 Ozone ISA.

General Comments

1. I compliment the United States Environmental Protection Agency for continuing to clearly characterize and communicate the effects of ozone as related to climate change building on the 2013 Ozone ISA to the current document, draft 2019 Ozone ISA.
2. Although evidence has increased supporting the relationship between tropospheric ozone and aspects of climate change, including a "causal relationship" with radiative forcing as well as a "likely to be causal relationship" with impacts on temperature, precipitation, and related climate variables, the causality determinations reached in the 2013 Ozone ISA are even further supported in the 2019 Ozone ISA, and I strongly concur with that position.

Further research would be useful, particularly quantifying the relationship between regional ozone RF including from ozone aerosols and other short-lived climate forcers on the hydrologic cycle, precipitation, and atmospheric circulation patterns; improving understanding of and ability to model critical ozone-climate processes; and continuing exploration of links between precursor pollutant control strategies, climate, and ozone concentrations. These research strategies would be extremely useful as we continue to better understand the role of ozone in the climate system scientific arena.

Dr. Sabine Lange

My main comment topics are discussed at the beginning of this document, with the details for each comment described after these summaries, followed by the expert consultant responses to my questions.

A reference list can be found at the bottom of this document for those studies that are not referenced in the ozone ISA.

Charge Questions: Please comment on the identification, evaluation and characterization of the available scientific evidence from epidemiologic, controlled human exposure, toxicological and associated human exposure and atmospheric sciences studies and the application of information from these studies to inform causality determinations for human health outcomes.

- Appendices 3 - 7 present assessments of the health effects associated with short-term and long-term exposure to ozone. The discussion is organized by exposure duration, broad health effects (e.g., asthma, ischemic heart disease, etc.), and scientific discipline. Please comment on the characterization of the evidence within these chapters.
- Please comment on the portrayal and discussion of the biological plausibility evidence presented in Appendices 3-7 and the extent to which: (1) the organization adequately captures the current state of the science with respect to potential pathways by which ozone could impart health effects, and (2) as currently constructed, inform causality determinations.

<u>Study Quality</u>

The EPA has improved their systematic review and study quality assessment. However, the study quality review needs further development.

- Only certain studies were included in the quality analysis – while the EPA notes that those from causal and likely causal designations, as well as those whose causal designations had changed were included, not all the studies from those appendices had quality evaluations in HAWC (for example, for the long-term metabolic epidemiology studies – only one of the 6 studies cited in causal designation Table 5-4 had a study quality evaluation listed in HAWC). There should be a more consistent inclusion of all relevant studies in the study quality evaluation.
- It is not clear how the study quality assessment is integrated into the main text and how it informs the conclusions based on the evidence.
- Given that the causality determinations explicitly state that there is evidence from a "high quality study" it is still not clear which studies the EPA considers to be "high quality". At the CASAC meeting on Dec 4, 2019 the EPA stated that all studies discussed in the review were high quality, but more information should be provided about how this decision is made, particularly given the fact that some studies that are discussed in the document are clearly higher quality than others.
- Chance, bias, and confounding are all potential reasons for a study to observe an association between two variables (Zaccai, 2004) and therefore should be more explicitly considered when presenting and discussing study results.

- More factors than just copollutants should be considered as important confounders in the referenced epidemiology studies.
- Results that are not statistically significant should be indicated as such in the ISA discussion. If there is a reason why statistical significance may not have been achieved (e.g., low sample size), this should be included in the discussion of the study results. The general conclusion from the expert consultants was that statistical significance does need to be given some consideration, but there are other factors such as patterns in the epidemiology data that also should be factored into the conclusions that are drawn.

Accuracy of Presentation

The EPA should provide a balanced summary of the study results for each health endpoint. Adequately communicating available positive, negative, and null results provides useful information for further documents in the ozone NAAQS review.

- In section summaries, divergent results should not be ignored, but rather should be included in a more nuanced summary of results. For example, the Arjomandi et al. 2018 study did not find an association between GSTM1 phenotype and inflammation. However, in the summary section for respiratory effects in healthy populations this divergent finding was not included or even intimated: the EPA stated that "Recent studies are consistent with previous findings and expand on observed interindividual variability in inflammatory responses, providing additional evidence that GSTM1-null individuals are more susceptible to ozone-related inflammatory responses."
- Further, information summarized from one section to another should maintain the accuracy and nuance of the underlying data. For example, in sections 4.1.16 and 6.2.4.1, the EPA states "Specifically, the evidence from controlled human exposure studies provided support for increased decrements in FEV1 and greater inflammatory responses to ozone in individuals with asthma than in healthy individuals without a history of asthma." This is not the case, the respiratory chapter addresses this point at length, and states that people with asthma are not less sensitive than people without asthma for lung function effects.
- The EPA should provide accurate study information as well as study conclusions. The study results presented in the metabolic chapter are particularly error-prone and need to be reviewed carefully. For example,
    - In section 5.1.4 (overweight and obesity) the EPA states that "Ozone exposure caused males on the control and high-fat diets to eat statistically significantly more food and trended toward statistically significant increases on high fructose diet (Gordon et al., 2016)." The exposure information is incorrect in the summary of this study (animals exposed one time per week, not 4 times per week). Also, the animals eat more, but they don't gain more weight - is ozone increasing their metabolism, such that they don't gain weight with increased caloric intake?
    - In section 5.1.5.1 (Other effects, inflammation), the EPA states that "Obesity-prone mice (adult male KK mice) were exposed to ozone for 13 consecutive weekdays [4 hours/day; Zhong et al. (2016)]." However, Table 5-10 says that the exposure was 3 consecutive days. The actual exposure in Zhong et al. (2016) was 13 weeks.
    - In section 5.1.5.4 (Other effects, serum lipids), the data presented for the Gordon et al. 2016 study is inaccurate. The EPA states that "The effect of high-fat and high-fructose diets was tested in male brown Norway rats" – the study was done in male and female

rats; “With ozone exposure (0.8 ppm ozone, 4 days/week for 3 weeks),” – exposure was 0.8 ppm ozone, 5 hrs/day, 1 day/week for 4 weeks (subacute exposure), or a single 0.8 ppm exposure for 5 hrs (acute); “Females were refractory to change.” – the abstract of the paper says “Female rats appeared to be more affected than males to O-3 regardless of diet.”

- In section 5.2.5 (metabolic syndrome and type II diabetes), the Jerrett et al. 2017 effect estimates are incorrect (presented are 1.28; 95% CI: 1.06, 1.55 and 1.20; 95% CI: 0.96, 1.50 with NO2 adjustment – should be 1.18 (1.04, 1.34) and 1.13 (0.97, 1.31) with NO2 adjustment.

Dose Assessment and Concordance

The EPA should appropriately compare animal to human ozone doses when extrapolating animal exposures to potential human risks.

- The PECOS statement for experimental studies in Appendix 3 (and on pp 3-19 and 3-26) notes that resting rats exposed to 2 ppm have an equivalent ozone deposition as exercising humans, citing Hatch et al. 1994. The EPA should further discuss that there is a direct relationship between resting human and resting rat ozone doses, and that a human with a ventilation rate that is five times higher than resting will have a 5 times higher dose. This should be correctly noted and Hatch et al. 2013 and McCant et al. 2018 (describes this misconception) should be cited.
- The EPA should also consider dose in their biological plausibility discussion, in particular the contrast between known personal exposure doses (which are typically very low because people spend most of their time indoors and ozone is largely an outdoor pollutant) and the concentrations at which the observed effects occurred.
- My summary from the opinions provided by the expert consultants on the questions of animal dosimetry is: Given that the causality determination for metabolic effects of ozone exposure is mostly derived from animal toxicological studies, it is appropriate for the EPA to more thoroughly discuss the dosimetric similarities and differences between animals and humans, beyond the simplistic reference to Hatch et al. 1994.

Clarity of Presentation

The EPA should clearly present the findings in each of the ISA sections, and should provide an accurate and balanced summary of results.

- When discussing the results from all studies, and particularly CHE studies, it is important to include the exposure duration (e.g. on pg 3-26 when discussing concentrations at which airway hyper-responsiveness has been observed) and the exercise level of the participants (e.g. in the integrated synthesis when discussing concentrations that could generate adverse effects in healthy adults).
- For the discussions addressing pre-existing conditions, the EPA should specifically note data that provides information on responsiveness of people with the condition to people without the condition (because this goes directly to potentially sensitive subpopulations). For example,

- In section 3.1.5.3 when the EPA discusses CHE studies, the EPA notes that in Horstman et al. 1995 they show that there are more wheeze symptoms in people with asthma exposed to ozone - is this compared to people without asthma, or ozone just increases wheeze in general?
- In section 3.1.6.2 the EPA describes the data showing whether there may be more sensitivity to respiratory effects of ozone of people who are obese or who have metabolic syndrome. However, in describing the study results, particularly of Ying 2016, Zhong, 2016, or Gordon 2016b, the EPA does not note whether there was a greater (or different) inflammatory response to ozone in the obese/metabolic syndrome animals versus lean/normal animals. Because this is the purpose of this section, these pieces of information should be included.
- The EPA includes sections about respiratory effects in pregnancy (3.2.4.7) and in populations with metabolic syndrome (3.2.4.8). Is the purpose of these sections to show that there is an increased response to ozone in these populations? If so, then the EPA should specifically provide information and discuss whether the data shows that these groups are more sensitive. As it stands, this conclusion is not clear.
- Is section 5.1.4 (overweight and obesity) intended to discuss the impact of ozone on obesity, or the effects of ozone on overweight/obese individuals? If it is the latter this should be part of the other chapters in the sections discussing sensitive subpopulations.

- If possible, the EPA should avoid making statements that address an unlikely conclusion, but that avoid addressing the conclusion of interest. For example:
  - In section 3.1.5.4 (lung function): "it was concluded that individuals with asthma were at least as sensitive to acute effects of ozone as healthy individuals." The conclusion of interest is whether people with asthma are *more* sensitive or not.
  - "However, despite limited evidence demonstrating increased sensitivity to ozone in individuals with asthma compared to those without asthma, there is consistent evidence that asthmatic individuals experience lung function decrements in response to acute ozone exposures." I don't think that there is a reasonable hypothesis that people with asthma would not experience ozone-induced lung function decrements.
- The EPA's underlying concern about people with asthma is perhaps not that they will have an increased innate response to ozone exposure (they do not seem to have greater lung function decrements, inflammation or airway hyperresponsiveness), but rather that they likely have less of a buffer against adverse effects. This is an important argument that EPA should emphasize when discussing the respiratory effects of ozone exposure on people with asthma.
- The EPA has described the exercise level in key CHE studies such as Schelegle et al. 2009 as a slow walking pace, but the authors of that study note that "This protocol contains six 50-minute exercise periods with minute ventilation maintained at 8 L/min/L of FVC (VE of approximately 40 L/min). As noted by Folinsbee and colleagues [Folinsbee et al. 1988] and McDonnell and colleagues [McDonnell et al. 1991], this level of exertion was ''intended to simulate work performed during a day of heavy to severe manual labor in outdoor laborers.'" This discrepancy in description should be clarified
- In section 4.1.8 (blood pressure changes and hypertension) ED visits and HAs, the EPA puts the findings into the context of the mean ozone concentrations measured in those areas. This is an attempt, I think, to understand the results in the context of a dose-response. This type of discussion is very helpful and should be included elsewhere.

- When the EPA states that there is little evidence for ozone impacting a particular endpoint, they should clarify whether there is little evidence because the studies haven't been done, or that the available studies do not show an association. For example, in section 4.2.2 (biological plausibility) the EPA notes that "However, considerable uncertainty remains in how long-term ozone exposure may lead to mortality given that there is little epidemiologic evidence of an association between long-term exposure to ozone and other cardiovascular endpoints such as IHD, stroke, or thromboembolic disease." Does this mean that studies have not been done, or that studies have been done that have not shown associations?
- In section 5.1.5.4 (Other effects, serum lipids), the EPA states in multiple locations that in an animal study, certain groups were "refractory to change". For example, "all other endpoints (HDL, LDL, and total cholesterol), ages, and doses (0.25 ppm ozone) were refractory to change (Bass et al., 2013)." Refractory means resistant and suggests that the endpoint would have changed but there was some active mechanism that kept them from changing. These endpoints weren't refractory, they just didn't change.

Consistency of Results & Reporting

The ISA would be strengthened by more justification of decisions in the face of conflicting evidence. An example of inconsistent (or seemingly inconsistent) results comes from section 5.2.3 (glucose and insulin homeostasis) where evidence is presented from three studies (Miller et al. 2016b, Gordon et al. 2013, Bass et al. 2013). Miller, Gordon, and Bass all came from essentially the same set of authors and tested effects of long-term ozone exposure in male rats. But they show different effects: Bass showed no change in fasting glucose with subchronic exposure, but Miller did; Miller showed decreased baseline insulin in subchronic exposed adult animals, but Gordon showed no change in adult exposed animals, and increases in insulin in senescent exposed animals. The EPA should speak to whether there are patterns in these results, or if the differences are spurious or show strain differences.

Consistency in EPA's reporting and interpretation of results is also important. For example, in section 4.1.17 (causality determination) the EPA states that "Studies from Europe, Canada and the U.S., several of which analyzed a large number of events per day in multiple cities, consistently reported null or only small positive effect estimates (i.e., OR $\leq$ 1.02) in analyses of MI, including for STEMI and NSTEMI (Section 4.1.5.1)." This is the only section where I have seen the magnitude of the association considered by EPA. Is there a reason why a small magnitude effect for this endpoint would be more important than a small magnitude effect for other endpoints? If the EPA is going to consider magnitude of effect for these studies, they should be clear as to why, and whether this is also a relevant consideration for other endpoints.

Applicability of Results from Animal Studies

Dose-responsiveness of effects of ozone exposure in experimental studies can be used to identify relevant biological plausibility pathways and exposure-specific responses, and so should be further discussed in those sections. In particular establishing no-effect and low-effect concentrations for endpoints such as long-term ozone exposure and lung function development would ease the extrapolation to effects in humans at ambient concentrations.

In addition, information about the comparability of animal models to human disease are useful in extrapolating results from animal studies – such as information about how good of a model allergic airway disease in mice is compared to humans. Even more important is information allowing the interpretation of *ex vivo* studies, such as experiments in isolated, perfused hearts (section 4.1.4).

Shape of the C-R Function

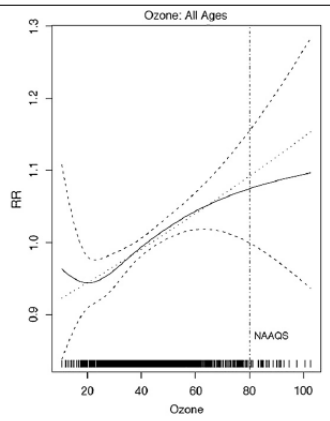
As was discussed in the CASAC's responses to the PM ISA and PA, errors and heterogeneity in epidemiology study variables can affect the apparent shape of the concentration-response (C-R) relationship and can obscure thresholds. Evidence for this has been provided by many peer-reviewed publications (Brauer et al., 2002; Cox, 2018; Lipfert and Wyzga, 1996; Rhomberg et al., 2011; Watt et al., 1995; Yoshimura, 1990) and notably by the EPA in the ISA preamble (US EPA 2015, Section 6c, pg. 29):

> "Various sources of variability and uncertainty, such as low data density in the lower concentration range, possible influence of exposure measurement error, and variability among individuals with respect to air pollution health effects, tend to smooth and "linearize" the concentration-response function and thus can obscure the existence of a threshold or nonlinear relationship. Because individual thresholds vary from person-to-person due to individual differences such as genetic differences or pre-existing disease conditions (and even can vary from one time to another for a given person), it can be difficult to demonstrate that a threshold exists in a population study. These sources of variability and uncertainty may explain why the available human data at ambient concentrations for some environmental pollutants (e.g., PM, O3, Pb, environmental tobacco smoke, radiation) do not exhibit population-level thresholds for cancer or noncancer health effects, even though likely mechanisms include nonlinear processes for some key events."

The problem described here is not whether a threshold in the data may exist, but rather that even if it does exist, epidemiology studies may not be capable of definitively identifying the threshold. To address this concern the EPA should explicitly acknowledge in the ozone ISA that variability and error in the variables can linearize C-R functions and obscure thresholds, and this acknowledgement should be included in those places where the EPA concludes that the relationship between ozone and a health effect is linear and has no threshold. I also recommend that the EPA begin to apply methods (and encourage the epidemiological community to apply methods) to address this particular concern, including errors-in-variables methods. If possible, the EPA should include these types of adjustments when applying the epidemiology C-R functions to their risk assessments.

In section 6.1.7 (shape of the C-R function), the EPA states that in the previous ISA the available studies showed no evidence of a deviation from linearity or the presence of a threshold for short-term ozone-mortality relationships. "However, it is important to note that the examination of the ozone-mortality C-R relationship is complicated by previously identified city-to-city and regional heterogeneity in ozone-mortality risk estimates (U.S. EPA, 2013a). Recent studies continue to provide evidence of a linear C-R relationship with no evidence of a threshold below which mortality effects do not occur along the distribution of ozone concentrations observed within the U.S." The EPA should provide information here noting whether the new studies address the consideration of city-to-city or regional heterogeneity that were concerns before, or if this is still an issue. If it is still an issue, the EPA should state it as such.

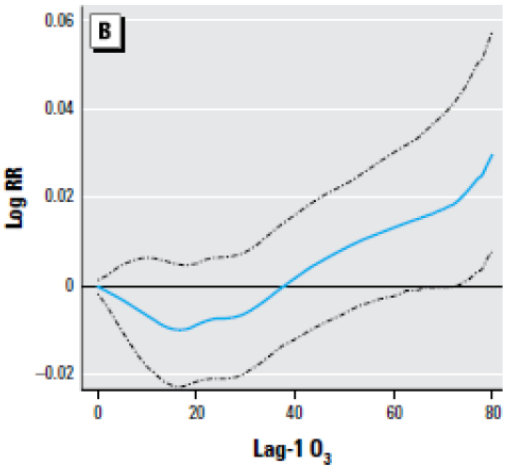
In addition, some of the plots that are presented by the EPA do not look linear and do appear to have a threshold, such as the Silverman plot (Figure 3-9), the Moolgavkar plot (Figure 6-6) and the Di plot (Figure 6-7). If the EPA thinks that there is so much uncertainty at the lower ends of these curves that we cannot trust the apparent U shape, then we also cannot trust that the shape is linear, and no conclusions should be drawn. For Silverman (Figure 3-9) there is still a lot of data at the point where the curve bottoms out (about 20 ppb), so the uncertainty cannot be all about data availability.



Note: The average of 0 day and 1 day lagged 8-hour daily max ozone was used in a two-pollutant model with $PM_{2.5}$ lag 0−1, adjusting for temporal trends, day of the week, and immediate and delayed weather effects. The solid lines are smoothed-fit data, with long broken lines indicating 95% confidence bands. The density of lines at the bottom of the figure indicates sample size. The NAAQS line indicated in the figure is reflective of a previous standard set in 1997. The form of this NAAQS was the 3-year avg of annual 4th highest daily max 8-hour concentrations.

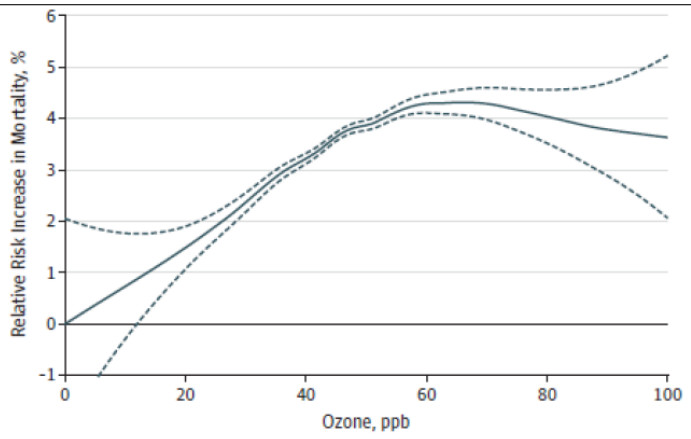
Source: Permission pending, Silverman and Ito (2010).

Figure 3-9 **Estimated relative risks (RRs) of asthma hospital admissions for 8-hour daily max ozone concentrations at lag 0-1 allowing for possible nonlinear relationships using natural splines.**



Source: Permission pending, Moolgavkar et al. (2013).

Figure 6-6 **Flexible concentration-response relationship for short-term ozone exposure and mortality at lag 1 for 24-hour avg ozone concentrations adjusted by size of the bootstrap sample (size of the bootstrap (*d*) = 4).**



Source: Permission pending, Di et al. (2017a).

Figure 6-7 **Percent increase in mortality for ozone in a two-pollutant model with $PM_{2.5}$ using penalized splines for both pollutants at lag 0−1 days in the warm season (April−September).**

Interpretation of Study Results

The ISA should address the adversity and clinical significance of the presented health effects. For example:

- In section 5.1.3 (glucose and insulin homeostasis) the EPA presents epidemiology results that demonstrate 0.5-3% chances in fasting glucose concentration. What is the clinical significance of this amount of change?

In addition, inclusion of the significance of relationships between different factors identified in epidemiology studies would help clarify the conclusions that can be drawn. For example:

- In section 3.2.4.1 the EPA discusses findings from the Children's Health Study noting that there is a relationship between ozone exposure, new onset asthma, and particular genetic variants. What is the significance of the relationship between genetic variants and new-onset asthma and ozone risk in the CHS study? What do those genes do? It is difficult to interpret genetic variant information in the absence of contextual information when determining the risk of new onset asthma with ozone exposure.
- In section 4.1.16 (effect modification) for pre-existing disease the EPA presents information about ozone-associated changes in HR or BP in people with mood disorders. What is the significance of mood disorders for ozone-associated changes in HR or BP?
- In section 5.1.5.1 (Other effects, inflammation), the EPA states that the Zhong et al. 2016 study showed an increase in inflammatory mediators in epididymal adipose. The EPA should provide information about the significance of increase inflammation in epididymal adipose, as well as whether inflammation was seen in other visceral adipose tissues.

It would be helpful if the EPA provided information about how short-term exposures impact long-term effects, how long-term exposures impact short term effects, and how effects of short- and long-term exposure are separated. For example:

- In section 4.1.8 (blood pressure changes and hypertension) the EPA notes that hypertension is a chronic condition that develops over a period of years. Therefore, how is an ozone exposure of several hours contributing to hypertension emergency department (ED) visits or hospital admissions (HAs)? Are the study authors (and the EPA) postulating that ozone concentrations trigger an immediate change in BP that sends the person with hypertension to the hospital? Sort of like an asthma attack? The EPA should clarify how ozone is expected to contribute to this endpoint.

Completeness of Study Information

The EPA should ensure that all relevant information is included in the study figures or tables. For example:

- In section 3.1.10.1 (copollutant confounding), the EPA notes that they provide study-specific details in the tables in Section 3.3. However, the information in those tables do not include the effect estimates for the copollutant models, only for the single pollutant models. The EPA should include the copolluant effect estimates in these tables, or in the text or figures of this section. The latter would be preferable, because of the importance of considering copollutant confounding. Similarly for the results that consider confounding by aeroallergens.
- Table 5-7 does not include all the information about the Ramot et al 2015 study – only one rat strain is included and not the 8 that were tested, and only one of the 3 ozone doses is included.

Causality Determinations

For the short-term ozone effects on metabolic endpoints, it is difficult to tell if the causality determination is warranted, because the presented data is not always accurate (see "accuracy in presentation" section of these comments). In addition, there is no consistent direction of effect presented – if biomarkers change in different directions in different experiments, does that matter for the EPA's causality determination? In addition, there is one CHE study presented that corroborates the animal studies, and one epidemiology study. However, multiple epidemiology studies are presented that have null results.

For the long-term ozone effects on metabolic endpoints there is again a problem of accuracy of reported results. Overall for this causality designation, there is limited epidemiology evidence (and that has issues, or associations are lost with copollutants, or copollutants aren't assessed, study quality was only assessed in one of the six epidemiology studies cited in Table 5-4, and at least one of the study quality details was wrong). The animal evidence is not always summarized correctly and shows somewhat inconsistent results. It does consistently show no effects at lower ozone concentrations (0.25 ppm), and all 3 of the cited animal studies were conducted in whole or in part by the same group of authors. Further analyses are provided in the detailed section of these comments.

For ozone and reproductive effects, the effects of ozone on male reproduction are based on little data (inconsistent epidemiology studies, one animal study), and additionally the EPA states for female reproduction that "In conclusion, results from epidemiologic studies are mixed, with benefits and detriments to female reproductive function with ozone exposures, while toxicological studies show limited evidence of effects on successful completion of pregnancy." Therefore, it is not clear why the EPA has designated effects on fertility and reproduction as "suggestive of causality".

Biological Plausibility

I encourage the EPA to present both positive and negative studies when presenting biological plausibility pathways, as well as information about exposure concentrations.

In section 5.1.2 (biological plausibility) the EPA needs to distinguish better between short-term and long-term effects of ozone on metabolism. It seems that they are ascribing short-term ozone exposures to diabetes development? In a comparable situation with asthma, the EPA does not attribute short-term ozone exposure to asthma development, but rather to exacerbation. The EPA states that "All of these upstream factors of autonomic activation and homeostatic imbalance can contribute to an animal model or humans being at a greater risk for developing metabolic syndrome or diabetes with ozone exposure." It is this last piece that is tricky to separate from chronic exposure effects. These axes are self-regulating, and although a single acute exposure may unbalance them, there is no evidence presented that this is irreversible. It seems like only unbalancing the system over and over again (chronic exposure) would predispose to metabolic disease.

Further Detailed Comments

Study Quality

- Based on the information in Appendix 10, the EPA has done a better job (compared to the PM ISA) describing the methods used for finding studies, screening them, and including or excluding them. I still have concerns about the study quality review, however. The EPA does note that there is a study quality review and they share the review criteria, and reference HAWC where the study quality assessments can be found. However, it is not clear how that information is integrated into the ISA, and I can't find the guidance text and prompting questions that EPA refers to on pg 10-21. In addition, the EPA states that they do not use individual study quality to assess results, but rather considers the quality of the literature as a whole. There are a few potential problems with this: 1) The causal determinations state that there is evidence from a "high-quality study" - without identifying what this study is, how can a reader tell if the causality determination is actually based on an objectively-determinable high quality study, or if in fact the literature as a whole has all of the pieces that could make a high quality study, even if any one study does not have all of those pieces? 2) The study quality does matter when drawing conclusions - there are certain kinds of conclusions that can only be drawn based on certain study types and quality, and if you are weighing contradicting evidence, then the weight should be placed more heavily on the higher quality study. If this is not made clear, then how does the reader know how to weight the studies (or how the studies were weighted by EPA?).

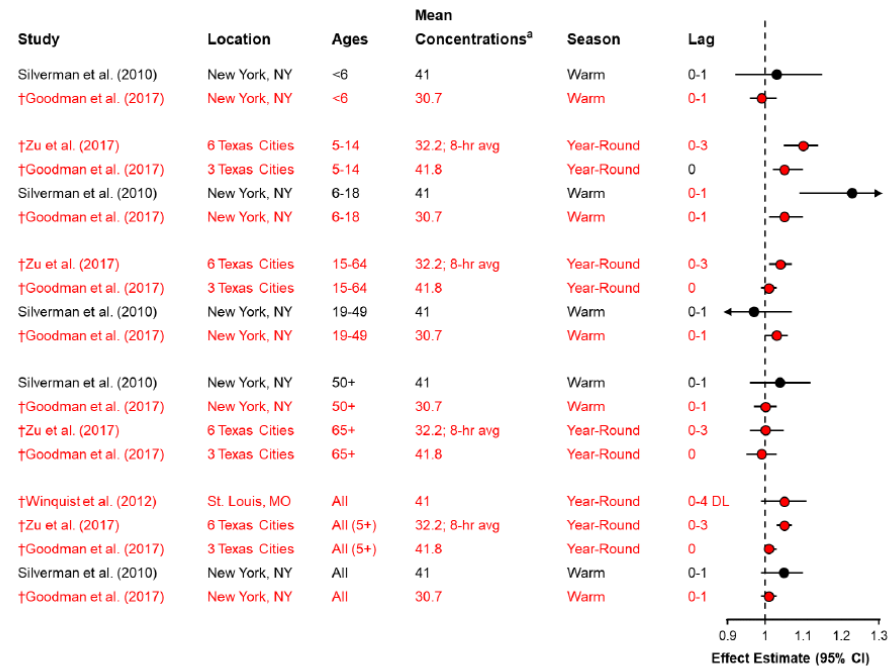
- It is not clear how the EPA chose to include studies in their HAWC study quality evaluation. For example, for section 5.2.5 (long-term ozone exposure and metabolic effects; metabolic syndrome and type II diabetes), only one of the three cited epidemiology studies had study quality evaluations in HAWC (Jerret et al. 2017, not Renzi et al. 2017, or Yang et al. 2018). For section 5.2.6 (metabolic syndrome and mortality) neither of the two cited epidemiology studies had study quality evaluations (Turner et al. 2016, Cruise et al. 2015).
- In the PECOS statement for experimental studies, the comparison group is stated to be the person themselves or an appropriate comparison group for CHE studies, or in animal tox studies to be a group exposed to a negative control. Negative controls are crucially important in these types of studies and should be required for both animal and human studies. In the inflammation section (3.1.4.3) the EPA notes the importance of the filtered air control by discussing exercise-induced effects, but then discusses results from studies without this control as if it is of the same value or quality as other studies that have the appropriate controls.
- In the integrated summary sections for respiratory effects, there is no distinction between those sections with lots of data (such as lung function), versus those with little data (such as symptoms in people with asthma - only one CHE study cited, with only single-city epidemiology studies and one panel study). Both are presented as having consistent evidence with no note about the strength of the overall database.
- In section 3.1.8, which discusses respiratory ED visits and HAs, study quality considerations are only discussed for one study, Winquist 2012, which did not show an association between ozone and all ages respiratory HAs (the EPA noted that only a single monitor was used for the exposure assessment). However, the same study was used without caveat as evidence for ED visits. The EPA needs to consider study quality for all studies, not just those that provide disparate information from the EPA's hypothesized effect.
- In section 6.1.6.1 (copollutant confounding), the EPA states that "The increase in the widths of the confidence intervals observed in these studies is consistent with a decrease in precision due to the limited data available to conduct copollutant analyses due to the PM sampling schedule." The EPA should provide information about how limited the data has really become. Often these NMMAPs studies have millions of data points, so dividing by 6 (for a one-in-six day PM sampling schedule) still provides hundreds of thousands of data points for the analysis. This would seem to provide adequate power for the analysis.
- In section 7.1 (reproductive and developmental effects), the EPA states that "Well-designed studies that consider sources of bias, including potential confounding by copollutant exposures, are emphasized." I do not see this statement made in the other chapters. Are the well-designed studies emphasized in other chapters as well?

Accuracy of Presentation

- In this document the EPA presents epidemiology study results with different averaging times on a single concentration scale, to allow direct comparison of results. However, using a simple concentration conversion does not capture the uncertainty that results when "converting" one averaging time to another, and may bias the resulting concentration estimates (Lange 2018). This has also been demonstrated by Anderson and Bell (2010), who found that interconversion amongst different averaging time metrics of ozone introduces uncertainty, and the ratios between the averaging times could differ across communities, as well as within communities by

temperature, season, and long-term ozone concentration. In addition, the Lange 2018 study shows that the 8-hr maximum and 24-hr average are not substantively correlated, so there should not be interconversion between the two.

- In section 3.1.4.4 the EPA notes that in the large and well-conducted MOSES study (Arjomandi et al. 2018) there is no association between GSTM1 phenotype and inflammation. However, this finding is diminished by the statement at the end of the section "Recent studies provide some further evidence that GSTM1-null individuals are more susceptible to ozone-related inflammatory responses, although the evidence is not entirely consistent." Then in the summary section for respiratory effects in healthy populations, this finding is further reduced to "Recent studies are consistent with previous findings and expand on observed interindividual variability in inflammatory responses, providing additional evidence that GSTM1-null individuals are more susceptible to ozone-related inflammatory responses.", where the discrepancy is not mentioned at all. A similar statement is made in section 3.1.11 describing the short-term ozone exposure respiratory effects causality conclusion. This sequence demonstrates my concern with how EPA summarizes evidence, by just dropping divergent results, and even in the face of conflicting evidence summarizes results as "consistent".
- In section 3.1.5.1 the EPA states that "Recent studies expand the existing evidence base and provide consistent evidence of an association between ozone and hospital admissions for asthma (Figure 3-4)." However, Figure 3-4 (reproduced below) shows mostly just a few positive associations for children 5-18, and a lot of null associations. This is not consistent evidence.
- In section 3.1.5.4 (lung function) the EPA presents data from Horstman et al. 1995 that summarizes the findings that FEV1 increases with individuals who use bronchodilators. Later in this section the EPA notes that Bertoli et al. 2013 had similar findings to Horstman et al. 1995, including that FEV1 decrements increased with a lack of inhaled corticosteroid treatment. This is the opposite result from Horstman 1995, not a consistent result.



DL = distributed lag.

Note: †Studies published since the 2013 Ozone ISA. Black text = studies included in the 2013 Ozone ISA.

[a]Mean concentrations reported in ppb and are for 8-hour daily max averaging times unless otherwise noted.

Results standardized to a 15-ppb increase in 24-hour avg, 20-ppb increase in 8-hour daily max, or 25-ppb increase in 1-hour daily max ozone concentrations. Corresponding quantitative results are reported in Table 3-5.

Figure 3-4 Summary of associations from studies of short-term ozone exposures and hospital admissions for asthma for a standardized increase in ozone concentrations.

- In section 3.1.6.3 describing respiratory responses in those with pre-existing cardiovascular disease, the EPA makes a fairly blanket statement that ozone enhances respiratory inflammation, pulmonary damage, etc, more in animals with hypertension/CVD disease compared to normal animals. But the summary of the studies shows that there is conflicting evidence of this - one whole set of studies shows lesser effects in the sensitive animals at lower concentrations of ozone, and another set shows that the age of the animal mediates the responses. More nuance needs to be applied to the summary of this information.
- In section 3.1.9 for respiratory mortality, the EPA concludes that there is a consistent positive association between ozone and respiratory mortality. Most of the evidence is from the 2013 ISA, with only a few studies adding to it - a multi city study showing an association, and a single-city study not showing an association. The evidence in the 2013 document is underwhelming - only about 1 page of discussion about 5 studies. Based on information from Table 6-5 of the 2013 ISA, my assessment of the presented studies is: Bell et al. 2005 was not statistically significant; the effect estimates in Katsouyanni et al. 2009 were only statistically significant in Canada in the summer (not the US or Europe in all-year or summer; or Canada in all-year); Klemm as noted in the ISA is negative, and Vanos et al. 2014 is statistically significantly positive (with much higher effect estimates than any other study). The Bell 2005 study was a meta-analysis and found non-statistically significant effects. This evidence is not consistent with the EPA's conclusions that there is a consistent positive association between ozone and respiratory mortality.

- In section 3.1.11 describing the causality conclusion for short-term ozone exposure on respiratory effects, the EPA notes that "The strongest single-day associations were generally observed with ozone concentrations on the same day as the outcome." However, according to the EPAs Table 3-1 this is not true, the strongest lags were distributed across the four days prior to the event, depending on the study.
- In section 4.1.11 (systemic inflammation and oxidative stress) the EPA concludes for panel studies that "Altogether, these epidemiologic panel studies provide evidence that short-term ozone exposure is associated with increased inflammatory responses." However, there is about an equal number of studies showing effects as no effects, which does not provide strong evidence for this conclusion.
- In section 4.1.16 (effect modification) for pre-existing disease, the EPA summarizes the respiratory effects evidence. It is not clear why this summary is included in the CVD section, it belongs in the respiratory section. In addition, the EPA states that "Specifically, the evidence from controlled human exposure studies provided support for increased decrements in FEV1 and greater inflammatory responses to ozone in individuals with asthma than in healthy individuals without a history of asthma." This is not the case, the document states that people with asthma are not less sensitive than people without asthma for lung function effects. This same mis-statement is also made in section 6.2.4.1 (long-term mortality, pre-existing disease).
- In section 4.1.17 (causality determination) the EPA states that "Animal toxicological studies published since the 2013 Ozone ISA provide generally consistent evidence for impaired cardiac function and endothelial dysfunction" However, the section describing animal toxicological studies of cardiac function (4.1.4.3) present as many studies/results showing no effect, as those that show an effect. This is not consistent evidence of impaired cardiac function.
- For appendix 5 (metabolic effects), because this is a new endpoint, it is particularly important for the EPA to provide balanced information about negative, null, and positive evidence.
- In section 5.1.3 (glucose and insulin homeostasis) the EPA states that "Recent epidemiologic studies provide some evidence of associations between short-term ozone exposure and these endpoints (Table 5-5)." However, the evidence presented is mostly null, or in opposite directions (one study showing increasing fasting glucose, another showing decreasing fasting glucose).
- In section 5.1.5.1 (Other effects, inflammation), the EPA states that "Obesity-prone mice (adult male KK mice) were exposed to ozone for 13 consecutive weekdays [4 hours/day; Zhong et al. (2016)]." However, Table 5-10 says that the exposure was 3 consecutive days. The EPA needs to carefully check all the study details in this appendix to ensure accurate reporting of results, as well as making sure that all of the details are complete (e.g. some study results present the number of hours per day for ozone exposure, and others do not).
- At the end of section 5.1.5.3 (Other effects, endocrine hormones), the EPA summarizes "Removal of the neuroendocrine input by surgically removing the adrenal glands removes the neuroendocrine stress activation, ameliorates the stress hormone response and attenuates glucose intolerance and other factors that contribute to metabolic syndrome in rodents exposed to ozone." As discussed by EPA in their study summary, adrenectomy only ameliorates some of the effects, so this is an overly general summary statement that does not correctly capture the data. It also does not capture the opposite effects of ozone on leptin in healthy versus diseased animals, which should be addressed.
- In section 5.1.5.4 (Other effects, serum lipids), the data presented for the Gordon et al. 2016 study is inaccurate. Here is the study summary that is presented on pg 5-20: "The effect of high-

fat and high-fructose diets was tested in male brown Norway rats. With ozone exposure (0.8 ppm ozone, 4 days/week for 3 weeks), there was significantly decreased serum cholesterol in animals on control diet, an effect which was ameliorated with high-fat or high-fructose diets (Gordon et al., 2016). In fact, ozone induced statistically significantly increased cholesterol in male animals on the high-fat diet versus high fat filter air controls. Serum triglycerides were significantly increased in ozone-exposed male rodents on the control or high-fat diets versus filter air controls. Females were refractory to change."

- "The effect of high-fat and high-fructose diets was tested in male brown Norway rats" – the study was done in male and female rats
- "With ozone exposure (0.8 ppm ozone, 4 days/week for 3 weeks)," – exposure was 0.8 ppm ozone, 5 hrs/day, 1 day/week for 4 weeks (subacute exposure), or a single 0.8 ppm exposure for 5 hrs (acute)
- "Females were refractory to change." – the abstract of the paper says "Female rats appeared to be more affected than males to O-3 regardless of diet."
- Actual results from the study:
  - Male Rats Acute: decrease in total cholesterol with ozone exposure in control group only; no change in glucose, HDL, or LDL; increase in triglycerides in control and high fat group but not in high fructose group with ozone exposure.
  - Male Rats Sub-acute: increase in HDL with ozone exposure in high fructose and high fat diets; no change in total cholesterol, glucose, LDL, or triglycerides.
  - Female Rats Acute: decrease in LDL cholesterol with ozone exposure in control group and high fat diets; no change in glucose, HDL, triglycerides, or total cholesterol
  - Female Rats Sub-acute: no change in total cholesterol, glucose, LDL, HDL, or triglycerides.

In section 5.2.3 (glucose and insulin homeostasis) evidence is presented from three studies. I have summarized the results from those studies below, having seen details that do not always agree with the EPA's summary:

- Bass et al. 2013:
  - Methods: 1-, 4-, 12-, or 24-month-old male Brown Norway rats were exposed to 0.25 or 1 ppm ozone for 6 h/day for 2 days (acute) or 6 h/day, 2 days/wk for 13 weeks (subchronic).
  - Acute exposure - all age groups (1, 4, 12, 24 mo) had increased glucose intolerance after 1 ppm ozone. Fasting blood glucose was elevated in 1-, 12-, and 24-month-old animals with 1 ppm ozone. 18 hrs after exposure glucose tolerance returned to baseline. No apparent effect of 0.25 ppm; increase in serum HDL at 1 ppm in 12-month-old rats only, no other changes in total or LDL cholesterol.
  - Subchronic exposure - all age groups had increased blood glucose after the glucose tolerance test compared to FA, but there was no increase in fasting blood glucose. No statistically significant effect of 0.25 ppm. Increase in serum HDL at 1 ppm in 12-month-old rats only, no other changes in total or LDL cholesterol. Significant decrease in serum triglycerides in 0.25 but not 1 ppm exposure in 24-month-old animals only.

- Miller et al. 2016b:
    - Male Wistar-Kyoto rats exposed to 0, 0.25, or 1 ppm ozone, 5 hrs per day, 3 consecutive days per week for 13 weeks.
    - Acute (1 week exposure) - fasting hyperglycemia and glucose intolerance at 1 ppm - reduced after 3 days of exposure compared to 1 day of exposure. No indication of impairment of peripheral insulin-mediated glucose clearance with ozone exposure. With 1 ppm ozone, no increase in insulin with glucose injection.
    - Subchronic - fasting hyperglycemia and glucose intolerance at 1 ppm - seems to be worse than after acute exposure. Completely reversed after 1 week of recovery. No indication of impairment of peripheral insulin-mediated glucose clearance with ozone exposure. With 1 ppm ozone, no increase in insulin with glucose injection. No change in liver or muscle insulin resistance.
- Gordon et al. 2013:
    - Male B/N rats exposed once per week 6 hrs per day for 17 weeks to 0 or 0.8 ppm ozone.
    - Serum insulin was significantly higher in old rats exposed to ozone, but not in 4-month old rats exposed to ozone. This is in contrast to the potential decrease in insulin seen in the Miller et al. 2016b study. Ozone had no overall effect on body weight.
- Miller, Gordon, and Bass all came from essentially the same set of authors. But they show different effects: Bass showed no change in fasting glucose with subchronic exposure, but Miller did; Miller showed decreased baseline insulin in subchronic exposed adult animals, but Gordon showed no change in adult exposed animals, and increases in insulin in senescent exposed animals. The EPA should speak to whether there are patterns in these results, or if the differences are spurious or show strain differences.
- In section 5.2.4 (adiposity, weight gain, and obesity), the EPA states that "Serum lipids (triglycerides and HDL cholesterol) were significantly changed with ozone exposure in aged animals [24-month-old males, 0.25 ppm ozone, 6 hours/day, 2 days/week for 13 weeks; Bass et al. (2013)], an effect not seen in younger animals with the same exposure."
- However, in the Bass 2013 study they present that: subchronic exposure to 1 ppm was associated with increased HDL cholesterol in 12-month-old animals only (not 1, 4, 24 months) nor at 0.25 ppm, and there were no other cholesterol effects. Table 1 of the study shows a decrease in serum triglycerides in 24-month-old animals with 0.25 ppm ozone but not 1 ppm, making this a suspect result (and this result is not discussed by the study authors in the text).
- In section 5.2.5 (metabolic syndrome and type II diabetes), the EPA states for the Jerret et al. 2017 study that "The study observed increased hazard ratios for incident diabetes (1.28; 95% CI: 1.06, 1.55); however, when adjusted for NO2, this relationship was slightly weaker and had wider confidence intervals (1.20; 95% CI: 0.96, 1.50)." As per the study, these numbers should be 1.18 (1.04, 1.34) (basic model + 10% criteria), and adding NO2 changes the estimate to 1.13 (0.97, 1.31). The estimates provided in section 5.2.9 discussing copollutant confounding, and the results presented in Table 5-12 are also incorrect. The HAWC exposure summary is also incorrect, noting for exposure that "The 2-year average ozone concentration prior to the diagnosis or last follow up was assigned to each study participant.". However, only the 2007-2008 concentrations of ozone from the CMAQ model were used in the study (although the follow up period was 1995-2011).
- In section 6.1.8 (causality determination), the EPA states that "Recent studies continue to assess the influence of potential confounders on the ozone-mortality relationship including copollutants,

temporal/seasonal trends, and weather covariates; overall, these studies report that associations remain relatively unchanged across the different approaches used to control for each confounder." There is actually quite a bit of variability in responses from different models with copollutants and temporal/seasonal trends (particularly the latter), which is not adequately summarized by saying that they are "relatively unchanged". Relative to what?

Dose Concordance

- To consider a rough estimate of the effect of exercise on inhaled dose, I multiplied the ozone concentration by the fold-increase in exercise ventilation for the amount of time that the subjects were engaged in exercise.
    - Controlled human exposure studies have shown significant respiratory symptoms in healthy humans exposed to 72 ppb for 6.6 hours with exercise (5 hrs at 40 L/min + 1.6 hrs at ~10 L/min, at 0.072 ppm, ~**1.6 ppm·hrs**)
    - Epidemiology study of asthma hospital admissions in Texas: statistically significant increase in asthma hospital admissions in 5-14 year-olds associated with 1 day (lag 0) 8-hr maximum ozone concentrations (average 8-hr max concentration of 41.8 ppb). Using some mean values, children spend about 10% of their time outdoors; personal ozone concentrations are ~20% of outdoor ozone concentrations. Children in these age groups will on average spend ~7.5 hours at a ventilation rate of 11.35 L/min (about double the at-rest ventilation rate). So, the average exposure is: 1 hr at 2X rest ventilation rate at 0.042 ppm+7 hrs at 2Xrest at 0.0084 ppm; ~**0.2 ppm·hrs**. Even considering a 2-fold difference in body surface area, there is a lot of dose difference amongst these studies.
- In section 3.2.3 (biological plausibility for long-term respiratory effects), the EPA does not provide the concentrations at which effects were observed. It is difficult to determine what effects are biologically plausible for occurrence in humans exposed to ambient concentrations of ozone (a typical annual average 8-hr max ozone concentration being in the range of 25-50 ppb), when experimental study results are presented without any concentration information. Concentration can substantially impact the effects of ozone (or any other chemical).

Clarity of Presentation

- The EPA should note when discussing the results of controlled human exposure (CHE studies) done at 60 and 70 ppb, what the actual measured exposures were, not just the target exposures (e.g. the actual exposure in the Schelegle et al. 2009 study was 72 ppb, not the target 70 ppb).
- Integrated Synthesis: "Under the assumption that respiratory symptoms might accompany similar ozone-induced FEV1 decrements, regardless of exposure duration, the model indicates that an 8-hour exposure to 64 ppb ozone concentration might reasonably be expected to cause an adverse response in young healthy adults." This statement mis-states "regardless of exposure duration", because the next statement notes that it is for "an 8-hour exposure". This discussion should also include a statement that this dose conversion assumes moderate to vigorous exertion for a substantial period of time.
- On page 3-26 the EPA describes pathways that researchers have demonstrated are activated in response to ozone to cause airway responsiveness. Are these pathways occurring simultaneously, or have different researchers found different pathways to be activated under different

circumstances? More information should be provided here. Also on page 3-26 the EPA notes that exposures of 0.8-2 ppm ozone cause airway hyperresponsiveness, but the exposure time of these should be included as well.

- In section 3.1.5.3 the EPA notes that in Horstman et al. 1995 the study finds that there are more wheeze symptoms in people with asthma exposed to ozone - is this compared to people without asthma, or ozone just increases wheeze in general? Also, the EPA states that "These observations are not changed by recently available studies or those in subsequent assessments (U.S. EPA, 12 2013a, 2006)." Does that statement mean that there are no other studies that have looked at this, or other studies have confirmed it?
- In section 3.2.1 (introduction to long-term ozone exposure and respiratory effects) the EPA presents the evidence from the last review as being very strong, and therefore does not make it clear why the causality determination was *likely* and not *causal*. This is the danger of emphasizing positive results. Some explanation of the likely causal choice should be provided.
- Section 3.2.4.5 addresses severity of respiratory disease. However, the evidence seems to be the same as just general associations with respiratory disease. For example, data is presented for HAs and ED visits, and for association with symptoms – this doesn't provide information about increased severity with ozone concentration. I suggest that the EPA further clarify why this data specifically provides severity information, rather than dividing the data up as is usually done (i.e. a section on symptoms, a section on ED and HA visits, etc). In addition, it is noted in the causality determination section that these endpoints are typically considered short-term and the studies do not usually control for short-term exposure effects. Given these concerns the EPA should be clear about how to interpret these study findings.
- In section 4.1.7.4 the EPA notes that "Importantly, this study also found that after 18-hours ozone exposure, both levels of ozone reduced increased sensitivity to aconitine-induced arrhythmia ($p < 0.05$). Similar results were also found in an another study by this group (Farraj et al., 2016)." This statement is very confusing and seems to suggest that ozone decreases the sensitivity of rats to arrythmia induced by aconitine. However, the Farraj et al. 2012 study states that "both low- and high-O3 exposure significantly reduced the total dose of aconitine necessary to elicit the first ventricular premature beat relative to air-exposed controls." This makes it clear that ozone increases sensitivity to aconitine-induced arrythmia. The EPA should clarify this statement.
- In section 4.1.16 (effect modification) the EPA states in their discussion of life stage effects that "Cakmak et al. (2014) used a population of 8,662 Ottawa and Gatineau patients referred for 24-hour ambulatory cardiac monitoring with exposure linked to the monitor closest to their home address. In subjects over the age of 50 ($n = 6,009$) cardiac rhythm was disrupted by an increased presence of heart block (1.13; 95% CI: 1.01, 1.27)." It is not clear what is meant by "cardiac rhythm was disrupted by an increased presence of heart block", and what that means in the context of ozone.
- In section 5.1.2 (biological plausibility) the EPA notes that "Diabetes and metabolic syndrome are disorders of the autonomic nervous system." This statement should be explained and referenced.
- In section 5.1.4 (overweight and obesity) the EPA states that "Genetically obese mice had airway hyper-responsiveness and responded more vigorously to acute ozone exposure than did lean controls (Shore, 2007)." Responded more vigorously in what way? To AHR, or some other response?

- In section 4.1.9 (heart rate and heart rate variability) the EPA notes that heart rate is a key prognostic indicator. They should state what heart rate is a key prognostic indicator *for*.
- In section 5.1.4 (overweight and obesity) in the summary of the Gordon 2017a study, the EPA notes that the purpose of the study was to "To determine the role of maternal exercise and diet on ozone's effect on glucose homeostasis and obesity in offspring, a study was conducted with multiple diet and exercise options." Then they note that while ozone exposure increased the glucose intolerance at one or two time points in every group compared to their matched FA controls, "comparisons were not made across groups." Since the purpose of the study was purportedly to determine how maternal exercise and diet affects ozone's effect on glucose homeostasis, it doesn't make sense that the study authors didn't test the difference between groups.
- In section 5.1.5.1 (Other effects, inflammation) it isn't clear whether the peripheral inflammation is only present in animals that are obese or have diabetes, or if this is generally associated with ozone exposure. This should be clarified.
- At the beginning of section 5.1.5.3 (Other effects, endocrine hormones), conclusions are presented about the effects of ozone on the endocrine system with citations to the review article by Snow et al. (2018). However, this section should be presenting the evidence for the presence or absence of these effects, not citing conclusions from a review article before presenting any data. I suggest moving this information to the end of the document, or removing it.
- In section 5.1.5.3 (Other effects, endocrine hormones), the EPA summarizes that "Elevated circulating stress hormones are consistently observed in animal models and in controlled human exposure studies after short-term ozone exposure." This should more precisely state that it has been observed in ONE CHE study.
- In section 5.1.5.4 (Other effects, serum lipids), the EPA states in reference to the Chen et al. (2016a) study, "No outcomes were reported for any metabolic endpoints evaluated with short-term increases of ozone exposure." If no outcomes were reported with short-term ozone exposure in the Chen 2016 study, then why is it covered here? Or does EPA mean that no metabolic outcomes were associated with short-term ozone exposure? If that is what EPA means, then they should say so.
- In section 5.2.2 (biological plausibility) there is a sentence exactly repeated more than once: "Ozone acts as a pulmonary irritant and stimulates nasopharyngeal and pulmonary nerves and receptors, including the trigeminal and vagal nerves, which induces downstream effects to the autonomic nervous system."
- In the first paragraph of section 5.2.11 (summary and causality determination) the EPA twice makes the statement "In prospective cohort studies in the U.S. and Europe increased incidence of type 2 diabetes was observed with long-term exposure to ozone."
- In section 6.2.3 (total non-accidental mortality) the EPA states "Generally, epidemiologic studies of long-term ozone exposure and total mortality use the 8-hour daily max ozone metric, though there are instances in some that use the 24-hour avg [e.g., Sesé et al. (2017)], or the 1-hour daily max [e.g., Jerrett et al. (2009)]." Presumably the EPA means here some long-term average of those different metrics is used in those studies?
- In section 6.2.4.1 (effect modification, pre-existing disease), the EPA seems to consider studies that do not compare people with the pre-existing disease, to people without the pre-existing disease (e.g. Zanobetti & Schwartz 2011 does not investigate people without a recent hospital admission, so they cannot conclude that people with a recent hospital admission have a higher

risk than those without). The EPA should consider studies with the appropriate comparators to differentiate whether a pre-existing disease makes a person more sensitive to ozone exposure.
- Section 7.1.3 (pregnancy and birth outcomes) should present effect estimates and confidence intervals with epidemiology study results, as is done in other parts of the document.
- Section 7.2 (nervous system effects) should present ozone concentrations used in the animal toxicology studies in the text, as is done in other parts of the document for these types of studies.
- In section 7.2.2.5 (neurodevelopmental effects), the EPA notes that "…in the current ISA these data are only reviewed in the long-term exposure section due to the sensitivity of the developing nervous system to toxicants and the potential for long-term impacts." This doesn't make sense - the divisions of short and long-term are made based on the exposure time, not the outcome. Mortality associated with short-term exposure could also be considered to have long-term impacts, but it is still included with short-term exposure. I agree with grouping together the neurodevelopmental effects, and I think that the long-term category is fine because in humans this can be up to a 9-month (i.e. not acute) exposure, but not because it may have long-term consequences.

<u>Consistency of Results</u>

- In section 3.1.5.2 the EPA notes that for asthma ED visits, that the evidence is consistent for both children and all ages - this is not what is typically seen for HA - is there a reason for this discrepancy?
- In section 3.1.7.2 the EPA notes that data from CHE studies suggest that ozone exposure activates the adaptive immune system, which may bolster defenses against infection. However, the epidemiology study results indicate an association between ozone and increased respiratory infections. The EPA should provide some discussion of how they reconcile these results.
- In section 5.1.5.4 (Other effects, serum lipids), the EPA summarizes that "Ozone exposure alters serum cholesterol in multiple animal models." However, the magnitude and direction of those alterations is not consistent. Some studies show increased HDL, some LDL, some no changes, and some CVD models show decreases in HDL or LDL. The EPA should discuss what these differential results mean in terms of the ozone mode of action and adversity of effects.
- In section 5.2.4 (adiposity, weight gain, and obesity), the EPA concludes from the epidemiology study results that "Recent evidence is limited, but provides some evidence that long-term exposure to ozone is associated with increased weight gain and obesity" However, the highest quality, most relevant to the US study showed no effect. In addition, the two Chinese studies that did potentially show effects did not have a study quality evaluation conducted by EPA - should the general conclusion be based on those studies?

<u>Applicability of Animal Study Results</u>

- The animal studies being cited in section 3.2.4.1 (new onset asthma) are being conducted at concentrations WAY HIGHER than ambient, and do not seem to be investigating a no effect level, which makes it difficult to know how comparable these effects are in humans. And even at these high concentrations effects weren't always consistent, or there was a potential dose response (e.g. Chou 2011 did not show increased BALF eosinophils with 5 biweekly 8-hr exposures to 500 ppb ozone, but Crowley 2017 did observe BALF eosinophils with 11 biweekly

exposures). Lee et al. 2011 demonstrated evidence of a threshold in rat studies of post-natal exposure, with less frequent exposures to 500 ppb ozone for 6 hrs per day having no developmental effects, compared to more frequent exposures.

- For the lung function and development section 3.2.4.2 the EPA presents animal studies of high exposures over multiple weeks causing lung function developmental effects. It is completely expected that course chronic exposure to that high a concentration of ozone is going to affect lung development, it has already been established that ozone can cause lung damage. It would be far more useful if a no and low-effect concentration was established. The rodent models do provide some useful information that there is a low concentration below which effects were not observed. The EPA should provide information about how this informs possible effects in human children.
- In section 4.1.4 (heart failure, impaired heart function) the EPA notes in the animal toxicological section that there were effects observed with ozone exposure in isolated, perfused hearts (McKintosh-Kastrinksy 2013). If this study involved direct exposure of the heart to ozone, how is this related to CVD effects that occur with pulmonary ozone exposure? The EPA should provide more information about how this type of study is relevant and how to interpret the exposure dose.
- In section 5.1.2 (biological plausibility) the EPA discusses the hyperthermic/hypothermic response in adult male rats caused by ozone. The EPA should provide more information about this to demonstrate its relevance to humans. For example, is this response also seen in other species/sexes? Is this thought to be comparable to humans? Is it dose-responsive?
- In section 5.1.2 (biological plausibility) the EPA notes the changes in thyroid hormone with ozone exposure in rodents. Given the known differences between circulating thyroid hormone and thyroid hormone responses on rodents compared to people, is there evidence that this mechanism would be relevant in humans?

<u>Interpretation of Study Results</u>

- In section 3.2.4.2 (lung function and development) the EPA notes that "To characterize lung health, lung function metrics capture the cumulative effects of pulmonary growth, damage, and repair (Wang et al., 1993). As such, measures of lung function are effective indicators of pulmonary effects related to exposure to environmental stressors." How are effects on lung function of acute exposures separated from effects of chronic exposures in these studies? That is, how do the authors know that they are measuring a long-term effect over time, and not just capturing the effects of the last day's exposure? The EPA should clarify these interpretations.
- The CVD plausibility section is supported by activation of the sympathetic nervous system, but in the respiratory section activation of the parasympathetic nervous system is cited. Can these occur simultaneously?
- In section 5.1.5.3 (Other effects, endocrine hormones), the EPA summarizes "In healthy rodent models, short-term ozone exposure was associated with either elevated serum leptin (Miller et al., 2015; Bass et al., 2013; Sun et al., 2013) or a trend toward increased leptin (Gordon et al., 2017b). In obese animals (Zhong et al., 2016) and diabetic animals (Ying et al., 2016), there was significantly decreased serum leptin with ozone exposure. Thus, healthy and diseased animal models have significantly different leptin responses to ozone exposure or the temporality differences between studies might explain the directionality differences." What is the potential

significance of these differences in leptin between healthy and obese/diabetic animals, and how can they be interpreted?

- In section 6.1.5.4 (effect modification, temperature), Figure 6-4 shows that there were some very high hourly ozone concentrations in this dataset (consistent with the data coming from 1987-2000) - are these patterns still relevant and observed at lower current ambient concentrations?
- In section 5.1.5.2 (Other effects, liver effects), the EPA presents metabolomic and gene and protein effects from several studies (Miller et al. 2015 and Theis et al. 2014). Were the results from these studies consistent with one another in terms of the pathways affected by ozone?
- In section 5.1.4 (overweight and obesity) the EPA states that "Studies done at the U.S. EPA examining effects of ozone at various concentrations (0.25, 0.5, 1.0 ppm) in healthy and obese rat models with leptin receptor mutation and associated cardiovascular disease demonstrated low sensitivity to ozone-induced lung injury and neutrophilic inflammation (Kodavanti, 2015)." I think that this is the wrong reference. Also, the statement suggests that animals with the leptin receptor mutation were less sensitive to ozone-induced lung injury and inflammation? Does this mean that EPA thinks that leptin is involved in the ozone-induced damage? Is this hypothesis in the respiratory chapter biological plausibility section? The EPA should consider separating out the results from these different rat models to be clear about what effect was seen in which model.

Completeness of Study Information

- At the beginning of section 5.1.5.2 (Other effects, liver effects) the EPA states that there is evidence for changes in the gut microbiome. However, there doesn't seem to be any direct evidence presented for that effect. Is this conclusion based on the change in bile acids? If so, that should be stated with appropriate caveats about the interpretation of this evidence, or the conclusion should be removed unless direct evidence is provided.
- Appendix 5 requires copy editing as well as careful attention to detail on exposure and experimental details.

Causality Determinations

For the epidemiology studies cited in this ISA, it seems that there is more work to be done on separating the effects of the many variables that can contribute to the measured endpoints. My question to the expert panel about the appropriate control for confounding, colliding, and mediating questions generated a general assessment that this kind of careful variable control maybe is not used frequently and remains an uncertainty in interpreting epidemiology study results.

Long-Term Ozone Exposure and Metabolic Effects (5.2). To justify the conclusion of "likely to be causal", the EPA states (in section 5.2.11):

- "Consistent epidemiologic evidence of increased risk of diabetes or metabolic syndrome" - based on Jerret et al. 2017 results from the Black Women's Health Study Cohort; Renzi et al. 2017 results from the Rome Longitudinal Study Cohort; and Yang et al. 2018 from the 33 Communities Chinese Health Study Cohort.
    - Only one of these studies was assessed for study quality by the EPA (Jerrett) and it had incorrect study details in the ISA and the HAWC study quality assessment. That study

showed a loss of significant association between ozone concentrations and incidence of type II diabetes when NO2 was included as a copollutant. Renzi et al. shows marginal associations with ozone at best (1.01 CI 1.00 to 1.02), and only in people under 50, but not older than 50. Yang is a Chinese study with substantial copollutant and other confounding concerns as noted by the EPA (these caveats were not included in section 5.2.11).
    - The EPA also cites results from Malmqvist et al. 2015 of effects on gestational or type I diabetes from developmental exposure to ozone (from section 7.1.3). However, this endpoint is not discussed in the causality designation, the study was not discussed in section 7.1.3, there was no statistically significant effect of ozone on gestational or type I diabetes, and there was no study quality evaluation.
- The EPA notes that the aforementioned findings are "consistent with two long-term ozone exposure studies in China, one in adults and one in children, presented increased odds of obesity in both adults and children as obesity is a risk factor for type 2 diabetes (T2D)."
    - These results are not consistent with results from the Black Women's Health Study Cohort (the same cohort as Jerret et al. 2017) that did not find an association between ozone and body weight gain (White et al. 2016). White et al. 2016 had a study quality evaluation in HAWC, while the two referenced Chinese studies (Dong et al. 2014, Li et al. 2015) did not.
- "Epidemiologic evidence of increased diabetes-associated mortality" – Based on diabetes mortality in the ACS cohort (Turner et al. 2016) and CanCHEC cohort (Crouse et al. 2016)
    - The EPA did not conduct a study quality evaluation of these studies. Why was that?
- There are a "limited number of epidemiology studies that evaluate potential copollutant confounding for PM or NOx" (and Jerrett et al. 2017 shows a loss of statistical significance of the effect when considering NOx confounding)
- "Animal toxicological studies of impaired glucose tolerance, fasting hyperglycemia, dyslipidemia, insulin resistance, and activation of the neuroendocrine pathway with ozone exposure"
    - The EPA summarizes the ozone concentrations associated with effects as 0.25 and 1 ppm, however, the available animal toxicology studies (Miller et al. 2016b, Bass et al. 2013, and Gordon et al. 2013) do not show effects at 0.25 ppm and Gordon et al. 2013 uses 0.8 ppm.
    - The animal evidence is not always summarized correctly, shows somewhat inconsistent results, and the 3 cited studies were conducted largely by the same group of authors.
- The EPA states that "The mortality findings are supported by epidemiologic and experimental studies reporting effects on glucose homeostasis and serum lipids, as well as other indicators of metabolic function (e.g., peripheral inflammation and neuroendocrine activation)." Studies of peripheral inflammation and neuroendocrine activation were not discussed in this chapter. The EPA also states that "In addition, these pathophysiological changes were often accompanied by increased inflammatory markers in peripheral tissues, and activation of the neuroendocrine system (Section 7.2.1.5)" However, the single study discussed in section 7.2.1.5 was a gene expression study after a single acute dose of ozone, with a return to baseline expression levels by 24-hours after exposure.
- The EPA states that "The animal toxicological studies provided evidence that long-term ozone exposure resulted in impaired insulin signaling, glucose intolerance, hyperglycemia, and insulin

resistance (Section 5.2.3.1). In addition, these pathophysiological changes were often accompanied by increased inflammatory markers in peripheral tissues, and activation of the neuroendocrine system (Section 7.2.1.5)" This statement lacks all the nuances of these studies, and the inconsistencies, and the fact that there was no indication of impairment of peripheral insulin-mediated glucose clearance with ozone exposure (Miller et al. 2016b).

- The EPA states that "Importantly, short-term ozone exposure studies also provided evidence that ozone exposure could contribute to the development of metabolic syndrome and show consistency with the evidence that long-term ozone exposure could lead to development or worsening of metabolic syndrome or its risk factors." However, there are several studies that investigated acute and subchronic exposure, and it was not clear that there were worse metabolic responses with increased exposure time (Miller et al. 2016b, Gordon et al. 2016b).

In section 7.1.2.2 (male reproduction), the EPA presents just a few epidemiology studies with somewhat conflicting results, and a single animal toxicology study. However, their conclusion is "Overall, there is evidence of impaired spermatogenesis and decreased sperm count and concentration from epidemiologic studies, and decremental effects on testicular morphology and impaired spermatogenesis from toxicological studies with ozone exposures." This is overly certain based on the available database.

Biological Plausibility

In section 7.1.2.1 (male and female reproduction and fertility, biological plausibility), the EPA does not make clear how the pathway leads from systemic inflammation to testicular and pregnancy effects. There is not connection provided there, and the "how" of the connection should be noted.

In section 7.1.3.1 (pregnancy and birth outcomes, biological plausibility), the EPA provides very little data to support the biological plausibility pathway. Further, there is not convincing evidence of systemic inflammation and oxidative stress induced by ozone exposure (a necessary upstream event in the pathway). The only statement supporting this upstream event is "The initial event of altered systemic oxidative stress is demonstrated in the epidemiologic literature with ozone-dependent increased odds of elevated CRP levels in nonpregnant individuals but CRP was unchanged at GD 5 in ozone exposed pregnant rodents (Miller et al., 2019)." There is no reference for the epidemiology study, and it provides much weaker evidence than the animal study.

Questions for Consultants

The EPA states in the ISA preamble that "Traditionally, statistical significance is used to a larger extent to evaluate the findings of controlled human exposure and animal toxicology studies. Understanding that statistical inferences may result in both false positives and false negatives, consideration is given to both trends in data and reproducibility of results. Thus, in drawing judgments regarding causality, the U.S. EPA emphasizes statistically significant findings from experimental studies, but does not limit its focus or consideration to statistically significant results in epidemiologic studies."

1) ***It has been established that associations found in an epidemiology study can be due to: causation, bias, chance, and/or confounding. If the concept of statistical significance is***

not useful in epidemiology studies, then how do the study authors/EPA rule out that chance has caused the observed association?

Responses:

Dr. Jaffe: "Regarding statistical significance in the preamble, I think the line right above your quoted line in the pre-amble is also important in this context "Other indicators of reliability such as the consistency and coherence of a body of studies as well as other confirming data may be used to justify reliance on the results of a body of epidemiologic studies, even if results in individual studies lack statistical significance." To me, what this is saying that if a group of studies, each with significance at say a 90% confidence all demonstrate a similar effect, then this can be used to justify a conclusion, even if no single study reaches a 95% confidence."

Dr. Jansen: "I was not aware of this. Imposing "significance" as a criterion on one type of study and not others seems wrong. On the other hand, does it explain the use of the phrase "positive associations" in many places (e.g., see page 4-2) rather than "significant positive associations?" I find such phrasing troublesome as it could be implying more rigor than exists and lacks clarity. If the result is null, it's null. Significant results should be shown distinctly, not diluted by "positive" null results, and weighted more heavily."

Dr. Lipfert: "Essentially, they do not. I take strong issue with the above EPA protocol on statistical significance and note that studies with wide confidence intervals are included. Experimental studies involve defined exposures with no co-pollutants or temperature effects and no measurement error. Variations among studies are thus real and should be analyzed as such. By contrast, epidemiological studies are subject to all of these sources of bias, especially temperature and the treatment of lags. Foley et al. (2003) reported that EPA considered outdoor ambient air quality from fixed ambient monitoring sites to be a "surrogate for exposure", which requires consideration of indoor exposures, by contrast with experimental studies."

Dr. North: "Statistical significance is useful in epidemiological studies, but in a limited way. These studies use regression to determine the association between a predictor x and a consequence y. This association might be statistically significant, that is, a good predictor, but causality could be absent: There may be a cause z that affects both x and y. Two examples: Children's shoe side predicts the children's reading ability. (Example due to Judea Pearl, in The Book of Why). Ice cream sales predict heat stroke cases. (I think Dr. Cox uses this one, which was more accurate before the era of air conditioning.) Progress for better prediction is to consider that there may be other factors that are predictive of y, get the data on these, and use these data on making the prediction. Children's age and high ambient temperature are candidates for the two examples."

Dr. Parrish: "I can respond to this question as a scientist with extensive experience in interpretation of results based on statistical significance, but not as an epidemiological expert. The concept of statistical significance is useful in interpreting the results of most scientific studies, but it is not of yes-or-no utility. Recent literature (e.g., Amrhein et al., 2019; Hurlbert et al., 2019) emphasizes the importance of not using statistical significance in a dichotomous manner, e.g. to decide whether the results of an analysis rules in or out any particular cause of an observed association. It is recommended to simply present p-values without label or category. With regard to the present question, the results of any

particular epidemiology study can provide an estimate of the probability that chance has caused the observed association, but that probability can never be reduced to zero. The same statement can be generally applied to the results of all studies that attempt to understand the cause of an observed correlation or association."

Dr. Thomas: "In addition to bias (of which confounding is one kind), chance can certainly lead to non-causal associations. Assessment of statistical significance is essential to judge the likelihood that an association could be due to chance, so it's incorrect to say that it "is not useful in epidemiology studies." Despite the longstanding and on-going debates about the usefulness specifically of p-values for this purpose (Greenland et al. 2016, Wasserstein and Lazar 2016), as opposed to a variety of other approaches (e.g., confidence intervals, Bayes Factors, etc.), they remain the most commonly used method for judging the possibility of chance. I do not see that the EPA has dismissed statistical significance testing in its evaluation of the evidence, although they correctly do incorporate "trends in data and reproducibility of results" as well as other considerations in their evaluation of the epidemiologic evidence."

Dr. Sax: "Epidemiological studies cannot be the sole basis for establishing a causal relationship because of the inherent limitations and because, in the case of observational epidemiological studies, you cannot rule out bias, chance and/or confounding with reasonable confidence. The issue is even more difficult when the observed effects are very small and not statistically significant or marginally significant. Because of these limitations it is essential to evaluate all lines of scientific evidence, including experimental evidence (human chamber studies, animal studies, mechanistic studies). By evaluating the consistency and coherence within and across the various scientific lines of evidence, one can obtain a better picture of whether a causal association is more or less likely. More importantly, the evidence may be able to also elucidate levels of exposure at which effects are more likely."

My summary of the responses: The general conclusion from the expert consultants was that statistical significance does need to be given some consideration, but there are also other factors such as patterns in the epidemiology data that also should be factored into the conclusions that are drawn.

Some short-term epidemiology studies use a method that is termed "case-crossover". These studies assess the pollutant concentration on the day of a health effect, and "control" days are those days when a person did not experience that health effect. My understanding is that the intention of this method is to control for intra-individual confounders. These study designs often use days before and after the health event (often matched to day of the week) as control days.

2) ***Am I correct in understanding that the intention of ozone case-crossover studies is to compare the ozone concentrations on a day when a health effect occurred for a person, to the ozone concentrations on a day when that health effect did not occur for that person?***
3) ***If so, then it would be important that some other factor (not related to ozone) did not prevent the health event from occurring on a control day. These studies often use days before and after the health event as control days, but for mortality studies (such as Di et al., 2017), how can a day after death be used as a control day? It doesn't matter what the ozone concentrations are after a person's death, that person would not be able to respond to that concentration. How should we interpret case-crossover studies that use control days after the event (particularly mortality) occurred?***

Responses:

Dr. Jansen: "2) I believe that is correct. I also seem to recall use of the ozone on, say, another Tuesday in a month, assuming the event happened on a Tuesday. 3) I believe your concern is not limited to mortality. I would expect using the day after for a hospital admission is affected by medical treatment and being confined indoors. If this is a fatal flaw, the studies should be excluded. If not, then there is clearly an uncertainty and the study results should be down weighted."

Dr. Lipfert: "2) Yes. 3) I am not an expert on case-control studies but I question their use in temporal rather than spatial studies. Ozone has both diurnal and seasonal trends as does ambient temperature, its primary confounder (see the Figures 2 and 3 above). Further, acute effects on mortality persist for several days, perhaps up to a week. Thus "case" days may not be independent of "control" days. Different people with different characteristics die on each day for different reasons. I would rather see both case and control periods extended for say, 3 days or more. The time-series model of Murray and colleagues (not cited in the ISA) considers temporal patterns of subject frailty, ambient temperature and air quality, each over several days. These temporal patterns are much easier to interpret than case-control findings. Di et al. combined spatial and temporal analyses, thus introducing geographic and climate variability (for no particular reason). I would have much rather seen a conventional time-series analysis in each of several locations involving socioeconomic differences as well."

Dr. North: "2) I'm not a fan of these cross-over studies, especially for mortality as the end point. For a good example, consider exposure to wildfires such as we have been experiencing in California. Consider hospital admissions for respiratory distress. What was the level the day before the smoke plume affected the area? What was the level the day the plume arrived? And after the wind blew the smoke away, then what was the level the next day? Yes, one might expect a low response level before the plume, and high levels after the high exposures, perhaps persisting for days after the levels have dropped. For Q3, the response is number of deaths on the exposed day versus the control day, and not the death of an individual person. (With cohort studies, it is more complex.) 3) I read the Di et al. study. Co-authors Schwartz and Zanobetti are among those trying to figure out how to do epidemiology where additional factors are considered. But I am not persuaded that confounding was not a significant issue for the results in the Di et al. study. The data base was all Medicare patients who died in a twelve year period. Most of the deaths occurred on days with ozone and PM2.5 levels well below the current standards. The death rates per 10 μg/m3 for PM2.5 was 1.45 per million persons at risk per day, and for 10 ppb ozone, 0.66 per million. These are extremely small numbers, but with sample size of nearly a hundred million days, the confidence limits were narrow around these numbers and did not include no increased risk. I looked up reference 9, Maclure (1991), on the study design. The Maclure Abstract begins, "A case-control design involving only cases may be used when brief exposure causes a transient change in risk of a rare acute-onset disease." I don't see the biological plausibility of comparing case days and control days for total mortality – not a rare acute–onset disease, but rather a situation where people who may be already very sick tend to die on days when they have additional stress. I suspect that high temperature may have acted as a confounding variable. Looking at figure 5 in the Di paper, I notice that the exposure response curve seems to flatten out (i.e., is insensitive to exposure level) for the higher 50% of the exposures, both for ozone and for PM2.5 . If these pollutants were causing the mortality increase, I would expect that the lower half of the exposure levels would be the flatter portion, and at higher exposures there would be more of a positive concentration response relationship. What may be going on is that in the days with higher half of the exposure levels, the pollutant levels are correlated (but rather

poorly) with the frequency of very hot days. On such very hot days mortality is significantly elevated. But on the lower pollution half of the days, there is a stronger correlation: a much lower frequency of very hot days. Very hot days can cause stress to an elderly person, especially in non-air conditioned space. Remember, the Di et al response rates are on the order of a one-in-a-million change. A small number of very hot days correlated with elevated exposure levels might give results such as reported in this paper."

Dr. Thomas: "2) Yes, that is the correct interpretation. An advantage of this design is that by making comparisons with an individual, between-individual confounding is completely eliminated, as are any factors that do not vary over time. While factors other than pollution that do vary over time, like weather, could still be confounders, these can be controlled in the analysis by standard statistical adjustment methods, as in case-control or time-series studies. 3) The original case-control design (Maclure 1991) involved a comparison of exposure at the time of the event (or some pre-specified time prior to it to allow for lag effects) to that at some previous comparable ("referent") time. For example, the referent time might involve the same day of the week to control for systematic weekly variation in pollution levels and/or confounders. My colleague, Bill Navidi (1998) pointed out, however, that seasonal variation and especially long-term trends in pollution levels could lead to bias if referent times always preceded event times, even if one or more entire year cycles were included; while there would be no bias if there were no long-term trends and if pollution followed a perfectly symmetric (e.g., sinusoidal) seasonal pattern, departures from such symmetry, as are common for both pollution and meteorology, would lead to bias. Instead, he proposed the "bidirectional case-crossover" design, in which two referent times, one before and one after, equally spaced around the event time, are used. The original Macluer design was intended to study personal time-varying characteristics such as behaviors that could be "triggers" for an event like death or heart attack; in this setting, it would be impossible to observe a behavior that occurred after death! In air pollution studies, however, personal behaviors are not being studied, but ambient exposures are and these can be measured and used meaningfully for comparison. While it is obviously true that pollution after the event could not be causally related to the event, the purpose of this design is to get an unbiased estimate of the expected exposure at the time of the event for comparison with the actual exposure at that time and can be interpreted as a sampling-based analog of the standard time-series approach for acute effects (Bateson and Schwartz 1999, Fung et al. 2003, Lu and Zeger 2007). Various versions of this design have subsequently been widely adopted in air pollution studies. Although the original bidirectional design has subsequently been shown to be slightly biased (Lumley and Levy 2000), a modified version involving using fixed time-strata, comparing exposures at event times within each stratum with those at all or selected times (e.g., day-of-week matched times) within the same stratum before and after the event, has been shown to be unbiased (Levy et al. 2001a, Janes et al. 2005a, Janes et al. 2005b), and this design has become the standard in substantive studies (e.g., (Levy et al. 2001b, Di et al. 2017)). As Mittleman (2005) says, "this strategy should be considered the de facto standard approach to the analysis of data arising in studies of the short-term effect of air pollution and weather" (see also references therein for additional studies using this design)."

Dr. Sax: "2) I believe that is the correct interpretation. 3) I think that this is a valid question, as you described by selecting days post-health effect this would violate an important epidemiological tenant for assessing a causal relationship – that is, that the exposure must precede the effect."

4) ***What is the importance of dose-concordance in establishing the biological likelihood of ozone-mediated effects occurring at relevant exposure concentrations in humans? Particularly in the context of known dose information about ozone: total inhaled dose includes concentration, exposure time, and exercise duration; Hatch et al., (2013) have shown that humans and rats that are exposed to ozone at rest achieve similar alveolar ozone doses, and that humans exercising at 5-times a resting ventilation rate achieved an ~ 5-times higher alveolar ozone dose; and that ozone concentrations are 2-10 times lower indoors where people spend most of their time.***

Responses:

Dr. Lipfert: "I fail to see any relevance. Controlled animal and human clinical studies serve only to show what might happen under controlled and idealized conditions. Null findings may thus be the most important. By contrast, epidemiology shows what actually does happen in the real world, including variability."

Dr. North: "That is a good question. My impression is that at ppm levels, humans and rodents are about equally sensitive, and that prolonged exposures at 5 ppm or higher are life-threatening to humans. There seems to be some information at lower levels. See my responses to Drs. Cox and Packham."

Dr. Thomas: "If by "dose-concordance" you mean comparability of doses to animals and humans from similar external concentrations and ventilation rates, I would expect that there are so many factors that differ that it would be unreasonable to expect the same dose-response relationships, even if doses could be scaled in comparable units."

Dr. Sax: "Particularly in the context of known dose information about ozone: total inhaled dose includes concentration, exposure time, and exercise duration; Hatch et al., (2013) have shown that humans and rats that are exposed to ozone at rest achieve similar alveolar ozone doses, and that humans exercising at 5-times a resting ventilation rate achieved an ~ 5-times higher alveolar ozone dose; and that ozone concentrations are 2-10 times lower indoors where people spend most of their time. I think this is a very important issue and one that has not been resolved or evaluated by EPA in weighing the evidence across different studies (i.e., human chamber studies and animal studies)."

5) ***Is there evidence that the animal models used to assess ozone effects (largely rats, mice, and non-human primates) are more, less, or similarly sensitive to ozone-mediated adverse effects compared to humans, at approximately equal inhaled doses?***

Responses:

Dr. Lipfert: "I don't know but, in my view, such tests should be only qualitative and used to study mechanisms. I don't see them useful to look for a "safe" dose since the actual human doses used in epidemiology remain unknown in part because of indoor effects."

Dr. Sax: "Again, this is a valid question, and given the evidence as presented in the ISA it is difficult to answer. I think that EPA's assertion that some of the high exposure levels used in the animal studies

(based on the Hatch et al., 2013 study) are relevant to ambient exposures in humans is likely to be simplistic at best, and a more detailed analysis to support an answer to this question is warranted."

My summary of the responses: Given that the causality determination for metabolic effects of ozone exposure is mostly derived from animal toxicological studies, it is appropriate for the EPA to more thoroughly discuss the dosimetric similarities and differences between animals and humans, beyond the simplistic reference to Hatch et al. 1994.

In this ISA I did not find population studies that considered causal pathways when assessing the association between ozone and health endpoints. It has been shown that the type of interaction between variables (e.g. confounding, colliding, mediating) can impact the results of regression analyses if these variables are controlled for in the regression equation.

6) ***In the absence of a causality diagram to direct the choice of variables to control in an epidemiological study, how can we judge whether a study has appropriately controlled for confounders, and has not inappropriately controlled for colliders (which can open up pathways between variables that otherwise would not be connected) or mediators (and thereby controlled away the effect)?***

Responses:

Dr. Jansen: "Clearly, the issue should be a key criteria used in the selection of and evaluation of studies. As I stated above, EPA needs clear criteria for study inclusion/exclusion, study quality, and causality classification. It is also not always clear which evidence is being given more weight than other evidence."

Dr. Lipfert: "I see causality and confounders in simpler terms: causality in terms of experiments and physiology and confounders in terms of bi-variate correlations and model evaluations with and without potential confounders. As stated above I believe that most of the mortality relationships are short-term and thus with few potential confounders. I don't think one can ever be sure that all of the long-term potential (spatial) confounders have been adequately controlled."

Dr. North: "Causality diagrams are still rare. But some epidemiology studies do consider multiple predictive factors, and explain how they do it. I expect we will consider this aspect in the risk assessment in the upcoming Ozone PA."

Dr. Thomas: "Very good question! Directed Acyclic Graphs (DAGs) can be useful tools for visualizing hypothetical relationships among observed and latent variables and for structuring an appropriate analysis strategy (Greenland et al. 1999). Investigators typically have such pictures in mind when conducting an analysis, although they are seldom presented formally in a substance matter publication (they are more commonly included in statistical methods papers). The basic principles that confounders must be controlled using the best available data on known risk factors (or surrogates for unmeasured factors in an attempt to minimize residual confounding), and that intermediate variables on a causal pathway from exposure to disease not be adjusted for, nor for colliders (that are determined by both exposure and disease but are not causal for disease), are well understood. The art is in deciding which variables are or are not appropriate to adjust for. While there are a variety of formal statistical methods

for dealing with adjustment uncertainty (Maldonado and Greenland 1993, Greenland 1996, Viallefont et al. 2001, Crainiceanu et al. 2007, Pope and Burnett 2007), it remains a matter for expert judgment, both by the original investigators and by critical readers.

Dr. Sax: "I don't think this is a new issue and as noted above, this is a particularly important limitation of observational air pollution studies. The study summaries that EPA presents in the ISA fall short of identifying the various limitations in the epidemiological literature and this remains an area of weakness in the overall evaluation of ozone health effects. For determining plausible (but not necessarily absolute) causation, a full integration of all lines of evidence is necessary. As noted previously, relying on only epidemiological evidence is not sufficient."

My summary of the responses: For the epidemiology studies cited in this ISA, it seems that there is more work to be done on separating the effects of the many variables that can contribute to the measured endpoints. My question to the expert panel about the appropriate control for confounding, colliding, and mediating questions generated a general assessment that this kind of careful variable control maybe is not used frequently and remains an uncertainty in interpreting epidemiology study results.

References

Anderson, G.B., Bell, M.L., 2010. Does one size fit all? The suitability of standard ozone exposure metric conversion ratios and implications for epidemiology. J. Expo. Sci. Environ. Epidemiol. 20, 2–11. https://doi.org/10.1038/jes.2008.69

Brauer, M., Brumm, J., Vedal, S., Petkau, A.J., 2002. Exposure misclassification and threshold concentrations in time series analyses of air pollution health effects. Risk Anal. 22, 1183–93.

Cox, L.A. (Tony), 2018. Effects of exposure estimation errors on estimated exposure-response relations for PM2.5. Environ. Res. 164, 636–646. https://doi.org/10.1016/j.envres.2018.03.038

Folinsbee, L.J., McDonnell, W.F., Horstman, D.H., 1988. Pulmonary function and symptom responses after 6.6-hour exposure to 0.12 ppm ozone with moderate exercise. J. Air Pollut. Control Assoc. 38, 28–35. https://doi.org/10.1080/08940630.1988.10466349

Hatch, G.E., Mckee, J., Brown, J., Mcdonnell, W., Seal, E., Soukup, J., Slade, R., Crissman, K., Devlin, R., 2013. Biomarkers of dose and effect of inhaled ozone in resting versus exercising human subjects: Comparison with resting rats. Biomark. Insights 8, 53–67. https://doi.org/10.4137/BMI.S11102

Lange, S.S., 2018. Comparing apples to oranges: Interpreting ozone concentrations from observational studies in the context of the United States ozone regulatory standard. Sci. Total Environ. 644, 1547–1556. https://doi.org/10.1016/j.scitotenv.2018.06.372

Lipfert, F.W., Wyzga, R., 1996. The effects of exposure error on environmental epidemiology., in: Proceedings of the Second Colloquium on Particulate Air Pollution and Human Mortality and Morbidity. pp. 295–302.

McCant, D., Lange, S., Haney, J., Honeycutt, M., 2017. The perpetuation of the misconception that rats receive a 3–5 times lower lung tissue dose than humans at the same ozone concentration. Inhal. Toxicol. 29. https://doi.org/10.1080/08958378.2017.1323982

McDonnell, W.F., Kehrl, H.R., Abdul-Salaam, S., Ives, P.J., Folinsbee, L.J., Devlin, R.B., O'Neil, J.J., Horstman, D.H., 1991. Respiratory response of humans exposed to low levels of ozone for 6.6 hours. Arch. Environ. Health 46, 145–50. https://doi.org/10.1080/00039896.1991.9937441

Rhomberg, L.R., Chandalia, J.K., Long, C.M., Goodman, J.E., 2011. Measurement error in environmental epidemiology and the shape of exposure-response curves. Crit. Rev. Toxicol. 41, 651–671. https://doi.org/10.3109/10408444.2011.563420

Watt, M., Godden, D., Cherrie, J., Seaton, A., 1995. Individual exposure to particulate air pollution and its relevance to thresholds for health effects: a study of traffic wardens. Occup. Environ. Med. 52, 790–2.

Yoshimura, I., 1990. The effect of measurement error on the dose-response curve. Environ. Health Perspect. 87, 173–8.

Zaccai, J.H., 2004. How to assess epidemiological studies. Postgrad. Med. J. 80, 140–147. https://doi.org/10.1136/pgmj.2003.012633

Dr. Corey Masuca

Appendix 1 – Atmospheric Source, Chemistry, Meteorology, Trends, and Background Ozone

1.3.1.2 Global and International Sources of Anthropogenic Ozone Precursors

While global and international sources of anthropogenic ozone precursors is adequately discussion, there appears to be a noticeable absence discussion(s) of localized, interstate and/or intercity anthropogenic ozone precursors.

1.5.1 Meteorological Effects on Ozone Concentrations at the Ground Level

There appears to be a noticeable absence of discussion(s) about topographical effects on/of ozone formation and ozone concentration transport (i.e., "bowl" inversion effects in valleys and areas surrounded by mountains).

Appendix 2 – Exposure to Ambient Ozone

2.4.2 Infiltration

While this section focuses on infiltration of outdoor ozone concentrations through/into indoors for home and buildings, have there been studies of ozone infiltration into vehicles, especially since most humans spend a significant amount of time in their vehicles during daily commutes?

Appendix B

Summary of Non-CASAC Member Consultant Responses to CASAC Questions

Following are CASAC questions (with some paraphrased explanations) sent to the non-CASAC member consultants, together with some responses received from them (not all consultants responded to all questions) and with some additional comments inserted. Emphases (bold) have been added throughout. Proposed one-word answers summarizing many of the responses (e.g., yes, no, or unclear) are provided, but the verbatim responses are frequently longer and more nuanced. CASAC member questions are presented in their entirety in Appendix C and non-CASAC member consultant responses are presented in their entirety in Appendix D.

Is the scientific information provided by the ISA clear? (*Does it provide testable predictions, expressed using unambiguous terms with stated operational definitions? Do the conclusions and generalizations presented follow from data and evidence using transparent reasoning and derivations? Are the implications of relevant scientific knowledge for effects of actions made clear?)*

Answer: No.

Explanation: Criteria for selecting and weighting studies, and the bases for drawing conclusions from them, are not clear. It is unclear how, if at all, the conclusions would change if consistent criteria were systematically applied for selecting, summarizing, and synthesizing studies. Causal determinations appear to be ambiguous and subjective: different people might make them in different ways, with no objective basis for determining from data which (if any) is right. What they mean or imply empirically is unclear. The treatment of wildfire contributions to ozone exposure, and implications for NAAQS, are unclear.

Responses from non-CASAC member consultants:

- Mr. Jansen: "I have always been frustrated by what I perceive as a **lack of clear criteria and transparency** in the descriptions of what leads to a particular causality classification. I can read several descriptions of evidence and am **unable to identify what makes one "causal" and another "likely" or "suggestive"** and as a result have a difficult time deciding whether I agree or not." "**I have always been concerned about the term causal vs. the term contribute.** We know death certificates are problematic with primary, secondary, and even tertiary causes listed. And I know from personal experience, **they are not necessarily correct**. **How does this factor into the analysis and messaging?**"
- Dr. North: "**Is the ISA *clear*? I do not find it so to me.**" **"It is not clear in the text whether these 'design values' include wildfire episodes**, or whether wildfire episodes with even higher ozone observations were removed as "natural disaster" exceptions." "The classification in **the causal categories are judgment calls. Different people might make them differently** based on the same supporting evidence." "I… have serious reservations about basing a risk assessment for health effects on the causal determinations from the framework EPA has used in this draft ISA. **I have difficulty interpreting what 'likely to be causal' means** in connection with possible confounding."

- Dr. Parrish: "In reading through Section 1.3 **a great deal of scientific information is summarized, but there is little or no discussion of the relevance of this science** to the NAAQS or the ozone design values upon which the NAAQS is based. **For example, wildfires** (here included in the broader category of landscape fires) and stratosphere-troposphere exchange are discussed. These are two natural sources of ozone that are specifically addressed in EPA's exceptional events rule. A crucial issue is the extent to which these sources can affect the ozone design value and perhaps cause an exceedance of the NAAQS that would be eligible for addressing under that exceptional events rule. How often do either of these natural sources cause exceedances? It is difficult to evaluate the significance of a particular scientific issue without the context of how that issue might affect the NAAQS that is being reviewed."
- Dr. Rhomberg: **"The treatment of "causation" in the ISA tends to be fuzzy** – fuzzier in some places more than in others – about the distinctions noted above. If patterns of association are plausibly explained by underlying causation, this is taken as sufficient evidence for such causation (when one should actually be comparing the hypothesized causative actions against other competing explanations for the patterns), and if such causation is inferred, it is taken to be universal, applying to other settings on the usually poorly stated (much less justified) presumption that the causation is universal and largely independent of other circumstances. The **measures of association are too easily taken as measures of the magnitude of cause, and these magnitudes applied to other settings without due consideration of how they may be contingent on differing circumstances**."
- Dr. Sax: "In my opinion, **the ISA could be clearer in many aspects**… Given the same information, I believe, that **different people (or different groups) could come to very different conclusions** and causal determinations based on current classification descriptions." "In addition, **the ISA causality framework would also benefit from some clarity regarding what each causal classification represents**. Specifically, is the classification meant to identify the strength of the evidence for a given level of exposure (e.g., below the current NAAQS) or is it simply to identify whether there is evidence enough of expected harm at any level of exposure?"

Is the scientific information provided by the ISA sound? Is it logically valid, correctly stated with appropriate caveats and based on the total body of relevant evidence?

Answer: No.

Explanation: The Draft ISA misinterprets evidence of association as evidence of causation without presenting data and analyses that rule out other interpretations (such as confounding, temporal trends, modeling choices, study selection biases, data aggregation biases, measurement error biases, etc.) It does not show that reducing ozone has caused or would cause any change in public health risks.

Responses from non-CASAC member consultants:

- Mr. Jansen: "I am of the school that a) **associations are not and cannot be causal** and b) quality human and animal experimental studies at relevant exposures need to be weighted over suggestive epidemiological (associational) studies to establish causality." (Tony Cox comment:

This school of thought is discussed further by Pearl 2009, 2010. It holds that causal and associational concepts "do not mix" because causation involves determining how changing one variable changes others, while association only describes whether and how observed levels of variables co-vary.)

- Dr. Lipfert: "**The ultimate test of causality is whether health has actually improved** since the late 1970s in response to peak O3 levels reduced by a factor of 5 in conjunction with coincident trends in spatial patterns of reduced smoking and improved medical care. **A search of PubMed found no support for such improvement.**"
- Dr. North: "**Is the ISA *sound*? Not in my judgment,** because there is little exploration of confounding and **too much reliance on strength of association** in the Bradford Hill criteria. I thought Appendix 2 [put] **insufficient emphasis on the differences between human exposure indoors, where most people spend most of their time, and ambient levels outdoors**. I found **weakness with the discussion of inflammation** in clinical studies at near-ambient exposure levels in Appendix 3, and I am **dissatisfied generally with the discussions of causality and confounding**."
- Dr. Parrish: "Section 1.3.1.3.3 **Landscape Fires deserves improvement.** It provides a synopsis of literature results, but there is no synthesis of the current state of knowledge. …**If wildfires have relatively large, episodic effects, why are these effects not seen in the ozone design values** in these rural northern U.S. states where wildfire impacts are expected to be most obvious?"
- Dr. Sax: "In general, **it is difficult to follow the rationale as presented in the ISA with regards to causal conclusion determinations** and **there are appear to be inconsistencies** in how the evidence is deemed sufficient to select one classification vs. another, such as in the cardiovascular vs. the metabolic outcomes. This highlights the need to have clear protocols and definitions regarding how these conclusions are derived."

Are the key scientific conclusions provided by the ISA scientific? Are they falsifiable by observations, tested against data, independently verifiable and verified, and logically sound?

Answer: No.

Explanation: The Draft ISA does not show the results of testing its conclusions and predictions against data not used in deriving them, i.e., predictive validity has not been established. As noted above, its key conclusions (causal determinations) appear to be subjective judgments expressed using ambiguous language: other people might make different judgments based on the same data, and there is no clear way to determine which judgment (if any) is right. Uncertainty about these vague, subjective judgments is not characterized.

Responses from non-CASAC member consultants:

- Dr. North: "Is the ISA *scientific*? The classification in **the causal categories are judgment calls. Different people might make them differently** based on the same supporting evidence."

- Dr. Sax: "The operational definitions of the causal **classifications leave a lot of room for subjective interpretation**."

Is the scientific information provided by the ISA policy-relevant? Does it quantify how changing ozone NAAQS would change (probable) public health outcomes, characterizing uncertainties, interactions, modifiers, and confounders?

Answer: No.

Explanation: The Draft ISA does not characterize whether or to what extent reducing ozone would cause (or has caused) lasting effects on public health risks.

Responses from non-CASAC member consultants:

- Mr. Jansen: "In addition to the issue of beneficial effects, there is the issue of recovery or reversibility. I did a search for the terms in the ISA and found some references to them in discussing experimental studies, which is appropriate. However, I did not see how it affected weighting nor causality classification. **In other words, if a metric was responsive but recovered, how is that evidence weighted and used in terms of causality classification?**"
- Dr. Lipfert: "My overall conclusion from the above is that **a risk-free NAAQS for ozone cannot be determined at any level above background** and certainly not as high as 70 ppb. It thus follows that **society must determine a tolerable level of risk, taking into account the presence of extreme frailty. This situation was not considered in the framing of the Clean Air Act**."
- Dr. North: "It seems to me an important issue **whether observed mild, apparently reversible effects** such as changes in FEV1 (forced expiratory volume in one second) seen **in healthy young exercising subjects imply a potential for adverse health effects in the general population**. What are the adverse health effects, and how well do FEV1 changes predict them? **What is the C-R relationship, not just for FEV1 changes, but for adverse health impacts that are persistent** and perhaps cumulative over time, such as scarring of lung tissue so that lung function is permanently lost?"
- Dr. Parrish: "In reading through Section 1.3 **a great deal of scientific information is summarized, but there is little or no discussion of the relevance of this science** to the NAAQS or the ozone design values upon which the NAAQS is based. For example, wildfires (here included in the broader category of landscape fires) and stratosphere-troposphere exchange are discussed. These are two natural sources of ozone that are specifically addressed in EPA's exceptional events rule. A crucial issue is the extent to which these sources can affect the ozone design value and perhaps cause an exceedance of the NAAQS that would be eligible for addressing under that exceptional events rule. How often do either of these natural sources cause exceedances? It is difficult to evaluate the significance of a particular scientific issue without the context of how that issue might affect the NAAQS that is being reviewed."

Can valid determinations of manipulative or interventional causation – that is, how and whether changing exposure would change health risks – be made based on observed associations of the types analyzed in the ISA?

Answer: No.

Explanation: Associations between past levels of variables do not predict how changing one variable would change another. Such predictions require causal analysis (Pearl 2009, 2010).

Responses from non-CASAC member consultants:

- Dr. Lipfert: "**No**." "**The ultimate test of causality is whether health has actually improved since the late 1970s in response to peak O3 levels reduced by a factor of 5** in conjunction with coincident trends in spatial patterns of reduced smoking and improved medical care. **A search of PubMed found no support for such improvement**."
- Dr. North: "I think this is a clear **NO**. CASAC should be seeking to evaluate manipulative or interventional causation, that is, determining how many people might be added or subtracted from having their health protected with an adequate margin of safety by a change in the primary NAAQS standard."
- Dr. Rhomberg: "**It is not possible to make an undisputable, totally sufficient conclusion of interventional causation,** because of the INUS nature of the possible causative processes." (Note: Dr. Rhomberg also explains that "The way these issues affect the 'causality' problem is that the complexities of context and potential interactions are what lead to the difficulty in inferring between the different aspects of causality – ability to affect, responsibility for observations, ability to generalize effect to other settings, ability to measure the impact of an effect, and ability to generalize that magnitude to other settings. They also affect the certainty with which causal determinations can be made, constituting the things that need to be thought through in characterizing uncertainty in any causation assertions." Dr. Rhomberg suggests that the NAS might provide useful advice on how methodological issues for causal determination framework can best be addressed.)
- Dr. Sax: "The short answer to this question I think is **no**."
- Dr. Thomas: "However, **the vast bulk of air pollution studies have not been designed or analyzed for the purpose of assessing manipulative or interventional causation**." *Note:*Dr. Thomas adds that "Nevertheless, the consistency of the findings from numerous observational studies, the concordance with short term human experimental studies (e.g., chamber or panel studies), and animal experiments, along with other lines of evidence supporting biological plausibility, as outlined in the preface to the ISA, allows a causal interpretation in terms of the likely effect of air pollution on the various health endpoints, if not a quantitative estimate of the predicted magnitude of the effect of a hypothetical intervention." However, the claimed consistency of associations is affected by modeling choices, study selection choices, averaging of disparate results, consistency in omission or incomplete control of confounders, and other factors that do not warrant a causal interpretation. Moreover, "a quantitative estimate of the predicted

magnitude of the effect of a hypothetical intervention" is then in fact based on these associations in the PA.

Is this actually a "formal causal framework"?

- Dr. Lipfert: "**No**, it's a list of subjective rationalizations based on studies selected from the literature according to unspecified procedures."
- Dr. North: "**No.** I agree that the terms are not clearly and unambiguously defined."
- Dr. Rhomberg: "**A formal causal framework" implies that a path to certain determination is available, and I do not think one is.**"
- Dr. Sax: "I am unsure of the answer/ or **answer is not clear**. I generally agree that the current **definitions for each classification leave a lot of room for subjective judgement** and in general it is not always clear (to me) how EPA weighs the evidence and comes to its final causal conclusions."
- Dr. Thomas: **"The approach used in the ISA does not exploit the emerging framework of "causal inference"** that constitutes one type of "formal causal framework." However, the "weight of evidence" machinery (Committee_to_Review_the_IRIS_Process 2014) used here is certainly a well-established and appropriate formal framework for reaching causal judgements combining evidence across scientific disciplines. The machinery of statistical causal inference is not capable of or intended to synthesize evidence across multiple studies from multiple scientific disciplines. … [For] a determination of a "causal relationship" the following is required: "… Generally, the determination is based on multiple studies conducted by multiple research groups, and evidence that is considered sufficient to infer a causal relationship is usually obtained from the joint consideration of many lines of evidence that reinforce each other." (Tony Cox notes: The WoE framework is not a formal or operational framework, insofar as it specifies no empirical operations or procedure for verifying or refuting a causal determination classification. Rather, the WoE franework offers a circular definition of causality: that a "causal" determination is generally based on "evidence that is considered sufficient to infer a causal relationship." This definition lacks verifiable empirical content. The WoE framework has not been validated for ozone. No studies showing that the WoE framework is "appropriate" (e.g., performing better than, or as well as, random guessing) have been published, as far as I have been able to determine by asking Chris Frey, Lianne Sheppard, and EPA. The causal inference framework that Dr. Thomas and others describe as "emerging" actually emerged approximately 100 years ago. It includes algorithms that are capable of, and intended to, synthesize evidence across multiple studies drawing on multiple scientific disciplines and both experimental and observational data. The misconception that formal causal analysis applies only to single data sets can be corrected by googling "integrative causal analysis" to obtain points of entry to the substantial technical on how to apply formal causal modeling and analysis algorithms to integration and synthesis of diverse data sets from multiple studies or fields.)

Does the ISA's causal determination framework clearly distinguish between necessary and sufficient causation?

Answer: No.

Responses from non-CASAC member consultants:

- Mr. Jansen: "**[T]he answer is no** simply because the framework does not accommodate those terms."
- Dr. Lipfert: "**No**."
- Dr. North: "**No.** Other factors should be considered, and positive association is not the same as necessary or sufficient. One needs to think in terms of partial causation through examination of multiple factors."
- Dr. Rhomberg: "**No,** as noted above, the naïve notion of causation sort of implies both necessity and sufficiency…"
- Dr. Sax: "The short answer to this is **no**."

Does a determination that exposure has a "causal relationship" with a health effect in a population imply that reducing exposure would reduce risk of the health effect in the population, other factors being held fixed? In other words, does a "causal relationship" determination imply a manipulative causal relationship?

Answer: No.

Explanation: Causal determinations in the Draft ISA are based primarily on associations and biological considerations. They do not address manipulative or interventional causation.

Responses from non-CASAC member consultants:

- Dr. Lipfert: "**Not in my opinion**."
- Dr. North: "**No.** I do not find that "causal" as used in the ISA implies "a manipulative causal relationship."
- Dr. Sax: "I think again the answer is **no**. … The observed associations do not imply causation…"
- Dr. Thomas: "That depends upon the context in which the term "causal relationship" is used. In the statistical literature on causal inference, yes, the goal is to estimate the effect of an intervention on differences in the expected outcome within an individual under different hypothetical scenarios. In the epidemiological literature -- and **as used in the ISA -- it refers to the existence of a mechanism under which exposure is a contributing factor**, which may imply that a change in exposure would be expected to change the outcome, but that is not the primary sense in which the term is used." (Tony Cox comments: "Existence of a mechanism under which exposure is a contributing factor" has no clear implications for whether or how much reducing exposure would reduce risk. Hence, it lacks clear policy relevance. Policy relevance requires addressing manipulative causation, i.e., how would a change in policy change

affect public health risks? Insofar as the ISA does not address or answer this question, it does not provide the most important scientific information needed to inform policy decisions.)

Can causal determinations be incorrect?

Answer: Yes.

Responses from non-CASAC member consultants:

- Mr. Jansen: "I have always been concerned about the term causal vs. the term contribute. We know death certificates are problematic with primary, secondary, and even tertiary causes listed. And I know from personal experience, **they are not necessarily correct**. **How does this factor into the analysis and messaging?**"
- Dr. North: **Yes**.
- Dr. Rhomberg: "**Yes, it can be incorrect.** It cannot be surely correct, so the question is how usefully to characterize the uncertainty and also how to act in the face of such uncertainty…"
- Dr. Sax: "I think that the causal classifications can be interpreted differently based on the current definitions and how EPA presents and interprets the data. That is, **they are open to interpretation**, particularly for determinations of a "causal relationship.""
- Dr. Thomas: "**Yes, of course, any human judgment could be incorrect.** That is true of those reached by a large body of experts as well but is much less likely!" (Tony Cox comment: Studies of groupthink, conformation bias, and group judgment have identified many factors that can make large bodies of experts even more likely than individuals to reach incorrect judgments. However, both individual and collective judgments about causality are notoriously error-prone and unreliable, making it preferable to rely on formal analytic procedures (e.g., conditional independence tests) applied to relevant data to draw causal conclusions.

I*s it clear how uncertainty about which category is correct should be (or has been) resolved* in assigning a final causal determination category, as in Table ES-1 p. ES-5) of the ISA?

Answer: No.

Responses from non-CASAC member consultants:

- Mr. Jansen: "I am a proponent of a quantitative uncertainty analysis being performed. And I would argue that **the ISA should include a section <u>in each chapter</u>on the literature that evaluates the uncertainty in the various components that will make up the risk assessment.** One could look to the outline provided in Dr. Lange's question number 4 (questions to consultants on the PM PA) for the beginnings of an outline, adapted of course for ozone. I do believe substantial data and hopefully studies exist to derive estimates for many of the items in her Table 1 and EPA should get on with performing that work. They have been advised to do so in the past. The approaches recommended by Dr. Anne Smith (see below taken from PM PA

comments of Dr. North) should also be tried and information embedded in the various studies needed for such analysis should be summarized."

- Dr. Lipfert: "**No**."
- Dr. North: **No.**
- Dr. Rhomberg: "I have argued that **this could be improved,** though it is hard to spell out a remedy…" (Tony Cox comment: One remedy might be to characterize the probability that each determination category applied to the available evidence. Assigning a single category when there is uncertainty about which is correct does not fully characterize relevant uncertainty for policy makers.)
- Dr. Sax: "**No, I don't think it is clear.** As noted above, EPA should provide more clarity and perhaps caveat the classifications based on the amount of uncertainty in the underlying evidence."
- Dr. Thomas: "While the process for deciding upon which category is appropriate is clearly described, **I do not see much if any discussion about [how] any disagreements about the choice of category were resolved,** except in terms of justification for changes in the categorization since the 2013 ISA (Tables IS-4 and IS-5)." (Tony Cox comment: My question was about uncertainty rather than about disagreements. Unanimous agreement on choice of category need not imply that there is no uncertainty; indeed, it might reflect a desire to conform when uncertainty is large enough to prevent easily defended disagreements.)

Is it clear how observations could be used to test and falsify a given causal determination if it is not correct?

Answer: No.

Responses from non-CASAC member consultants:

- Dr. Lipfert: "Laboratory experiments may determine whether a given type of response *can* happen. Epidemiology is required to determine under what circumstances it *does* happen."
- Dr. North: **No.**
- Dr. Rhomberg: "There are methods to try to uncover contingencies on other factors, by measuring them and using directed graph analysis."
- Dr. Sax: "**No,** I don't think that as the framework is laid out by EPA, it could be applied consistently and that the same causal determinations would be necessarily developed by different groups. As noted above, this is exemplified by the change in classification for two key health outcomes (CV and mortality) in the current ISA."
- Dr. Thomas: "While it is always possible that new data will emerge that leads one to question a previous determination, such speculation would be beyond the scope of the ISA." (Tony Cox comment: My question was whether causal determinations are falsifiable by data (including existing data), and if so, how. For example, if particulate matter in a carefully performed accountability study were permanently reduced by 30 or more micrograms per cubic meter by a

ban on coal burning, and if this reduction had no observed effect on mortality rates on time scales of years to decades in multiple relevant comparison groups, would this outcome falsify a causal determination of "causal"? Why or why not?)

Is the correctness of each causal determination in table ES-1 formally and transparently evaluated in the ISA? In other words, have formal rules for determining the correctness of the causal determinations in Table ES-1 (p. ES-5) from the data and evidence presented been explicitly stated, applied systematically, and the results documented? (If so, where?)

Answer: No.

Responses from non-CASAC member consultants:

- Dr. Lipfert: "**No.** I find Table ES-1 useless and completely subjective."
- Dr. North: "**No.** I do not believe the determination represented by the entries in Table ES-1 can be defended as correct. I think the Administrator, CASAC, and we who are advising CASAC could argue *ad nausium* about causality, **and we would be better off trying to address how to estimate the extent of health response** and the severity of the health response for exposures in the 60-70 ppb range."
- Dr. Rhomberg: "I have argued **that a more rigorous conception of causation is indeed needed**. Bradford Hill considerations and other such approaches are at best guides for fuller thinking through, they are not tests to be passed."
- Dr. Sax: "**No,** see my comments regarding the NAAQS framework above, as well as Goodman et al. (2013) reference (above)."

Does a determination that an exposure-response (or concentration-response (C-R)) relationship is a "causal relationship" imply that it is <u>entirely</u> causal, with no contribution from incompletely controlled confounding, modeling errors and biases, or other non-causal sources?

Answer: No.

Responses from non-CASAC member consultants:

- Dr. North: "**No.**"
- Dr. Rhomberg: "**No."**
- Dr. Sax: "This is an interesting question, and I think that the way that EPA uses the C-R functions in the risk assessment, assumes that there is a causal relationship between the air pollutant (in this case ozone) and the various health effects, under the conditions specified in the underlying epidemiology model and then EPA often applies this to other populations. However, **the epidemiological studies are flawed in that they do not account for uncontrolled or incompletely controlled confounders**, **and have errors and biases** that are not always discussed or caveated. Therefore, it is more likely that there does exist some lower bound of the attributable fraction and it is unclear how much of it is necessary for consideration of a true causal effect."

- Dr. Thomas: "**No, but that depends upon what is meant** by "entirely causal." Epidemiologists have long recognized a "complex web of causation (MacMahon and Pugh 1970) meaning that no single factor is ever both necessary and sufficient to cause disease. A "causal relationship" is generally held to mean that a risk factor is a real component of one of the "sufficient component causes" of disease (Rothman 1976)."

Does a determination that a C-R relationship is a "causal relationship" imply 100% certainty that it is causal?

Answer: Unclear.

Responses from non-CASAC member consultants:

- Dr. North: "**No.** Let's use probabilities instead of seeking yes-no answers."
- Dr. Rhomberg: "**No**"
- Dr. Sax: "I think as EPA uses it, **yes**."
- Dr. Thomas: "Obviously **that depends** upon the confidence with which that judgment has been reached." (Tony Cox note: The confidences with which causal determination judgments in the ISA have been reached are not specified.)

Does a determination that a C-R relationship is a "causal relationship" imply that it is causal for every member of a population, or might it be deemed "causal" if it is causal for a sensitive subpopulation only?

Answer: No.

Responses from non-CASAC member consultants:

- Dr. North: "**No.** Sensitive subpopulations must be considered."
- Dr. Rhomberg: "**No**"
- Dr. Sax: "EPA currently uses C-R functions in its risk assessment based on selected epidemiological studies. The epidemiological studies are based on different population groups (sometimes children, sometimes the elderly, sometimes all ages etc.), therefore I think **it depends on the epidemiological stud**y and the population group that the study includes."
- Dr. Thomas: "**Not necessarily**."

Are the five categories mutually exclusive?

Answer: Unclear.

Responses from non-CASAC member consultants:

- Dr. North: "I must respond that the lines between the categories are **unclear to me**. I cannot attest to mutually exclusivity. Again, I think the level of ozone exposure is critical. CASAC should evaluate the significance to health (and welfare) of ozone exposures in the 60-100 ppb

range, in the context of other factors. Causality for adverse health effects seems well established at higher levels in the occupational health literature, for example, the Ozone Material Safety Data Sheets. See https://www.cdc.gov/niosh/idlh/10028156html; www.amsbio.com/images/featureareas/ozilla/Ozilla-MSDS.pdf."

- Dr. Rhomberg: "**Unclear**, partly because the meaning of each category could be tweaked. What is clear is that the categories do not seem well suited for making the distinctions between existence of some causal property, the degree of independence of that property from other circumstances, and the possibilities and limits of generalization."
- Dr. Sax: **"I don't think they are necessarily mutually exclusive**. Again, as defined I think the scientific evidence could be evaluated differently by different people or groups such that they would arrive at different conclusions (or different categories) given the same evidence. See for example our assessment of the cardiovascular effects of ozone and EPA's conclusions in the prior ozone ISA. Prueitt, RL; Lynch, HN; Zu, K; Sax, SN; Venditti, FJ; Goodman, JE. 2014. "Weight-of-evidence Evaluation of Long-term Ozone Exposure and Cardiovascular Effects." Crit. Rev. Toxicol. 44(9):791-822. Goodman, JE; Prueitt, RL; Sax, SN; Lynch, HN; Zu, K; Lemay, JC; King, JM; Venditti, FJ. 2014. "Weight-of-evidence Evaluation of Short-term Ozone Exposure and Cardiovascular Effects." Crit. Rev. Toxicol. 44(9):725-790."
- Dr. Thomas: "**Yes.**"

Are the five categories collectively exhaustive?

Answer: Unclear.

Responses from non-CASAC member consultants:

- Dr. North: "I am not sure I care about "collectively exhaustive." I do not think the framework is useful."
- Dr. Rhomberg: **Unclear**
- Dr. Sax: "**Not necessarily**, as there is no counter classification to each of the levels"
- Dr. Thomas: "**Yes.**"

Can a body of evidence be categorized as "likely to be causal" if the probability of causality based on the evidence is less than 50%?

Answer: Unclear.

Responses from non-CASAC member consultants:

- Dr. North: "I prefer probability statements and partial causation with multiple factors as in Dr. Cox's example. I do not like implied "bright lines," such as greater than 50% means likely. Such conventions need to be agreed to among the users. I do not believe that is the case here."
- Dr. Sax: "Based on EPA's definitions, it is **hard to say** what percent probability of causality needs to be for the evidence to be considered "likely to be causal.""

- Dr. Thomas: "Causal inference methods aim to estimate the "Average Causal Effect", not the probability of causality.The "Probability of Causation" (PC) is an estimate of the probability that a specific individual's disease was caused by some aspect of his exposure history, essentially an individualized version of the epidemiologic concept Population Attributable Risk Fraction." (Tony Cox comment: My question was about probability of causality for a body of evidence in a population, and what level justifies a label of "likely to be causal," not about probability of causation for an individual.)

Is it clear that the ISA's study selection process has successfully provided a comprehensive, trustworthy, and unbiased selection of the best available science on ozone and health effects? Is it clear why results from Moore et al. (2008) are included and cited as "key evidence" but contrary results from Moore et al. (2013) are excluded? More generally, *is it clear that study inclusion and exclusion criteria were applied systematically and neutrally to identify and select the best and most up-to-date studies to inform the ISA's conclusions?*

Answer: No.

Responses from non-CASAC member consultants:

- Mr. Jansen: "**Clear criteria are needed for study inclusion/exclusion**, study quality, and causality classification. It is also **not always clear which evidence is being given more weight** than other evidence." "[In the past], when significant negative and null studies were included there was a tendency for the narrative to critique (dismiss?) them and simply accept the significant positive studies. **This suggests a bias in the review of the evidence.** Identified missing studies also ran the risk of post-rationalization. EPA needs to work harder to critique all studies, weight them appropriately, and avoid post-rationalization of additional studies identified in the review process. Unfortunately, this bias is continuing."
- Dr. Lipfert: "As evidence of **subjective selection of epidemiology studies**, I note that only one of my studies is included (see Table 1 above), but the authors saw fit to mention a study of O3 exposure and erectile disfunction. …Studies with wide confidence intervals were included."
- Dr. North: "I and others have found **other relevant studies not included**. There is a deficit of published papers on interpretation of the available studies, especially on uncertainty, variability and severity of health effects. Discussion of inflammation is an example. Non-experts have difficulty understanding the importance of the range of biomarkers for assessing the degree to which these biomarkers indicate a public health impact deserving of protection under the language of the Clean Air Act. I did not find any references to the journal *Risk Analysis*, for which I am an area editor, and which publishes many papers on air pollution health risks. I expect other journals may be neglected as well."
- Dr. Parrish: "Another significant shortcoming in Section 1.3 is that the **uncertainty of the ozone precursor emissions estimates should be clearly discussed and defined** to the extent possible."

- Dr. Sax: "In my opinion, **the ISA could be clearer in many aspects, including how studies are selected for inclusion (or exclusion)**, and **how the evidence is weighed** and specifically used to reach causal conclusions. …**The ISA could benefit from being clearer on how the literature is selected (i.e., more specifics on search criteria used, screening on literature, literature included and excluded, and reasons for inclusion and exclusion).** This would provide greater transparency related to this process."
- Dr. Thomas: "I agree that the **Moore (2013) should be discussed**."

Are there other studies that are omitted from the ISA that should be included?

Answer: Yes.

Responses from non-CASAC member consultants:

- Dr. Lipfert: "I summarized ozone mortality risk estimates for both long-and short-term exposures from my publications and from the literature in Table 1 (note that **my publications have not been included in EPA ISAs**. In general, there is a lot of heterogeneity among these estimates, even within the same cohort."
- Dr. North: "I am concerned that **EPA's selection process is leaving out studies with negative findings for ozone**. This is evidence against the ISA study selection process being comprehensive, trustworthy, and unbiased." "I would like to see possible interventions made explicit at the regional level and evaluated in terms of **how much these interventions would reduce exposure, and then, what impacts these would have on health and welfare measures**, with results in probabilistic form. **Anne Smith's 2019 paper in *Risk Analysis*** demonstrates how such an evaluation might be done." "I did not find **Kai [Chen] et al. (2018) in** the references in ISA Appendix 3, and that concerns me a great deal [https://journals.lww.com/environepidem/FullText/2018/03000/Does_temperature_confounding_control_influence_the.2.aspx]."
- Dr. Parrish: "[T]here is evidence (Oikawa et al., 2015; Almaraz et al., 2018) **that NOx emissions from agricultural soils plays a major role in degraded air quality in the Imperial Valley, located in an ozone no-attainment area in California. This issue deserves more discussion** since it may play a significant role in an important no-attainment areas in the country."
- Dr. Sax: "**[S]everal studies** (some of which I am an author of and references provided throughout) **are missing from the ISA. More transparency is needed on how EPA selected studies for inclusion/exclusion**." "I refer Dr. Frampton to several studies that were published by me and my colleagues, and the findings from these evaluations are more consistent with the change in the causality determination for CV disease. **None of these studies were mentioned or included in the ozone ISA**: Prueitt, RL; Lynch, HN; Zu, K; Sax, SN; Venditti, FJ; Goodman, JE. 2014. "Weight-of-evidence Evaluation of Long-term Ozone Exposure and Cardiovascular Effects." Crit. Rev. Toxicol. 44(9):791-822. Goodman, JE; Prueitt, RL; Sax, SN; Lynch, HN; Zu, K; Lemay, JC; King, JM; Venditti, FJ. 2014. "Weight-of-evidence Evaluation of Short-term

Ozone Exposure and Cardiovascular Effects." Crit. Rev. Toxicol. 44(9):725-790. Goodman, JE; Prueitt, RL; Sax, SN; Pizzurro, DM; Lynch, HN; Zu, K; Venditti, FJ. 2015. "Ozone Exposure and Systemic Biomarkers: Evaluation of Evidence for Adverse Cardiovascular Health Impacts." Crit. Rev. Toxicol. 45(5):412-452. Petito Boyce, C; Goodman, JE; Sax, SN; Loftus, CT. 2015. "Providing Perspective for Interpreting Cardiovascular Mortality Risks Associated with Ozone Exposures." Reg. Tox. Pharmacol. 72(1):107-116."

- Dr. Thomas: "I do not have the comprehensive knowledge of the air pollution epidemiology literature to address whether there are key studies missing, other new evidence, or inadequately considered weaknesses."

Are there studies included in the ISA that should be omitted (e.g., because of uncontrolled confounding, obsolete or incorrect modeling assumptions, conclusions dependent on unverified assumptions, ecological fallacy, lack of causally relevant information, lack of design that can support valid causal inferences, or other methodological problems?)

- Dr. Sax: "Other than general study quality criteria that EPA has identified in an Annex to each of the Appendices, it is still unclear how EPA weighs studies based on study quality**. This is an area that EPA still needs to work on and add to its ISA evaluations**. After an assessment of study quality, which would help to identify methodological issues or biases in certain studies, EPA **could develop criteria for excluding studies** or at the very least for giving certain studies less weight in the overall conclusions and causal determinations."

Is it clear that the process followed in selecting and summarizing scientific studies in the ISA was sufficient to assure accurate, unbiased, up-to-date, and trustworthy summaries of the relevant scientific literature to inform causal determination judgments?

Answer: No.

Explanation: Neither the selection of studies nor the summary of studies that have been selected appears to be unbiased, comprehensive, or trustworthy. For example, I searched PUBMED for "ozone respiratory effects causal." Three of the top seven articles returned are as follow, shown here with selected conclusions (emphases added):

- Qian Z, He Q, Lin HM, Kong L, Zhou D, Liang S, Zhu Z, Liao D, Liu W, Bentley CM, Dan J, Wang B, Yang N, Xu S, Gong J, Wei H, Sun H, Qin Z; HEI Health Review Committee. Part 2. Association of daily mortality with ambient air pollution, and effect modification by extremely high temperature in Wuhan, China. Res Rep Health Eff Inst. 2010 Nov;(154):91-217. "Among the gaseous pollutants, we also observed statistically significant associations of mortality with NO, and SO2, and that the estimated effects of these two pollutants were stronger than the PM10 effects. The patterns of NO2 and SO2 associations were similar to those of PM10 in terms of sex, age, and linearity. ***O3 was not associated with mortality***."
- Cox LA Jr, Popken DA. Has reducing fine particulate matter and ozone caused reduced mortality rates in the United States? Ann Epidemiol. 2015 Mar;25(3):162-73. doi: 10.1016/j.annepidem.2014.11.006. "***There were no significant positive associations between***

changes in PM2.5 or O3 levels and corresponding changes in disease mortality rates between 2000 and 2010, nor for shorter time intervals of 1 to 3 years."

- Goodman JE, Prueitt RL, Chandalia J, Sax SN. Evaluation of adverse human lung function effects in controlled ozone exposure studies. J Appl Toxicol. 2014 May;34(5):516-24. doi: 10.1002/jat.2905. "Overall, ***these studies do not demonstrate a causal association between ozone concentrations in the range of the current National Ambient Air Quality Standard and adverse effects on lung function***."

None of these negative results is mentioned in the draft ISA.

For studies that that *are* cited in the ISA, I performed the following spot checks and found the following results.

- Page 3-91 of the ISA states that "A limited number of recent studies provide evidence of an association between long-term exposure to ozone and asthma development in children. … An overview of the evidence is provided below. A recent CHS analysis examined asthma incidence in relation to improved air quality in nine southern California communities (Garcia et al., 2019). ***Decreases in baseline ozone concentrations in three CHS cohorts, enrolled in 1993, 1996, and 2006, were associated with decreased asthma incidence.***" However, Garcia et al. (2019) actually state that "Among children in Southern California, decreases in ambient nitrogen dioxide and PM2.5 between 1993 and 2014 were significantly associated with lower asthma incidence. ***There were no statistically significant associations for ozone*** or PM10." (Garcia E, Berhane KT, Islam T, McConnell R, Urman R, Chen Z, Gilliland FD. Association of Changes in Air Quality With Incident Asthma in Children in California, 1993-2014. JAMA. 2019 May 21;321(19):1906-1915. doi: 10.1001/jama.2019.5357. Emphasis added.)
- Table 3-3 on "Summary of evidence for a likely to be causal relationship between long-term ozone exposure and respiratory effects" cites the study of Moore et al. (2008) ("Ambient ozone concentrations cause increased hospitalizations for asthma in children: An 18-year study in Southern California") as providing "key evidence" for the ISA's causal determination that there is "a likely to be causal relationship between long-term ozone exposure and respiratory effects." Specifically, Moore et al. is cited as providing "***Consistent evidence of an association between long-term ozone concentrations and hospital admissions*** and ED visits for asthma." Yet, follow-up work by Moore et al. (2012) noted methodological limitations of the 2008 paper (especially, that its results may have resulted from incorrect untested modeling assumptions, rather than from information in the data) and provided and applied an improved methodology ("CMRIER" or "causal models for realistic individualized exposure rules"). A key result was that the previous significant effect of ozone was no longer found. (Moore et al. (2012) state that "The results from the original HRMSM analysis based on the continuous ozone variable estimated with the G-computation method resulted in an estimate of an increase of 1.44e-06 in the proportion of asthma-related hospital discharges for a one-unit increase in ozone. [This is the 2008 study cited in Table 3-3 of the ISA.] ***Unlike results from the HRMSM analysis with the continuous ozone variable, the CMRIER results are not significant.*** Note that the HRMSM analysis was based on G-computation estimation which ***artificially relies on untestable parametric modeling assumptions*** to estimate HRMSM parameters when the ETA

assumption is violated. Thus, in this ozone study [the 2008 study cited by the ISA], ***significant results from the G-computation analysis may be a consequence of the approach taken and not solely based on the information in the data***." (Moore KL, Neugebauer R, van der Laan MJ, Tager IB. Causal inference in epidemiological studies with strong confounding. Stat Med. 2012 Jun 15;31(13):1380-404. doi: 10.1002/sim.4469.) This more recent paper is not mentioned in the ISA. The ISA cites the 2008 results as "key evidence" without noting that the authors subsequently revised them in the 2012 paper.

Responses from non-CASAC member consultants:

- Dr. North: "I am concerned that **EPA's selection process is leaving out studies with negative findings for ozone**. This is evidence against the ISA study selection process being comprehensive, trustworthy, and unbiased. Including Moore (2008) but not Moore (2012) is further evidence of weakness…"

Do you find in the Executive Summary a clear explanation of the extent to which the key evidence supporting the ISA's causal determinations consists of, is sensitive to, or is derived from unverified modeling assumptions, or from modeling assumptions that more recent literature has found to be incorrect or inadequate? Have you found information in the ISA on sensitivity of causal determination conclusions to untested, uncertain, or incorrect assumptions? (If so, where? See Table Annex 6-1, cf p. 6-67 for a discussion of what should be done. Has it be done, and is it clear what the results were?)

Responses from non-CASAC member consultants:

- Dr. North: "I am disappointed that the Executive Summary focuses on the causality determination, rather than a common sense summary of the relevant science including discussion of modeling assumptions, uncertainties and variability in ozone exposure. I did not find the Annex to Appendix 6 useful as a guide to "what should be done." It seemed like a defense of EPA practice as best practice, and I disagree, especially on the use of the Bradford Hill criteria."

Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" adequately controlled for potential confounding and residual confounding by variables such as income and weather variables?

Answer: No.

Responses from non-CASAC member consultants:

- Dr. North: "Issues of **how to deal with confounding by temperature, extent of air conditioning in homes and workplaces, and socioeconomic status – interrelated factors that will differ by location – need to be carefully evaluated in order to get good estimates of mortality and morbidity responses. EPA in this ozone ISA seems far behind evolving 'best practices'** in how to do such analysis."
- Dr. Thomas: "Of those studies cited in support of these determinations that I am familiar with, the authors have gone to appropriate lengths to control for such confounders, to the extent

possible with the available data. **Of course, residual confounding can never be excluded** from any observational epidemiology study." (Tony Cox comment: Residual confounding by variables such as income and temperature can easily be excluded by using continuous values instead of categorizing them.)

Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" have adequately controlled for biases due to exposure estimation errors or exposure misclassification errors?

Answer: No.

Responses from non-CASAC member consultants:

- Dr. North: "Estimating exposure 'according to the centroid of the postal code' is quite crude, and misses all the subtleties whether children are outdoors, indoors in air conditioned space, or indoors in non-air conditioned space, and the extent to which they are exercising in these environments. It also may miss exposure to materials the aggravate asthma, such as pet and cockroach dander."
- Dr. Thomas: "While the various studies differ in their methods of exposure estimation, **few to my knowledge have used formal methods of measurement error correction.** However, the bias from measurement error would generally be in the direction of reducing effect sizes and power, not introducing false positives." (Tony Cox comment: In multivariate regression models, measurement error can, and often does, introduce false positives and increase estimated effects sizes.)

More generally, is it clear how criteria for individual study quality were applied to each study used in making causal determinations, and what the results were? (See Table Annex 6-1, cf p. 6-67.) *Is it clear how the limitations of each individual study were taken into account in causally interpreting their reported associations and in making causal determinations?*

Answer: No.

Responses from non-CASAC member consultants:

- Dr. North: "**No**, I find the discussions of study quality rather superficial. It is clear that some of the recent studies have initiatives for investigating confounding factors. Much more is needed, in the context of learning how to do better estimation of the C-R relationship in the low exposure range."

Does the ISA make clear how its causal determinations would change if evidence from associations caused by confounding, residual confounding, measurement error, or unverified modeling assumptions were excluded?

Responses from non-CASAC member consultants:

- Dr. North: "**I did not find much discussion on confounding, measurement error, and unverified modeling assumptions.** The discussions I did find were often superficial."
- Dr. Thomas: "This seems rather speculative, absent any evidence that those studies included have failed to adequately address these possible biases." (Tony Cox comment: There is abundant evidence, including discussions by the original authors in papers such as that of Tétreault et al., that included studies did not adequately address these and other possible biases.)

Is the biological evidence presented in the ISA to support causal determinations correctly stated, correctly interpreted, relevant for predicting effects of changes in the ozone NAAQS, and up-to-date?

Responses from non-CASAC member consultants:

- Dr. North: "I would like to see more evidence in the ISA of the kind in the Michaudel et al. and Xu et al papers, indicating mechanisms for health damage and biomarkers or indicators for biological changes in humans at much lower exposure levels. …It will be important to get expert judgment on which cytokines, neutrophils and other indicators of inflammation are most significant for predicting irreversible damage to lung tissue such as described in the Michaudel et al. quote above. The **support in the ISA seems weak for inflammation in humans at 60-80 ppb exposure over 6.6 hours with exercise. This support is from studies done before the last review of the ozone standard**."
- Dr. Thomas: Yes.

Does the biological evidence presented in the ISA provide well-validated scientific information suitable for predicting the effects on public health of changing NAAQS standard for ozone?

Answer: No, the ISA does not address the effects on public health of changing NAAQS standard for ozone. (Tony Cox comment: Modeling the effects of changes in the PA without first providing a critical discussion of the scientific basis for such models in the ISA deprives readers of the scientific information needed to make an informed decision about whether changes in policy are needed to protect human health.)

Responses from non-CASAC member consultants:

- Dr. North: "Yes, there is some useful information, and I will consider a series of published papers reporting similar findings as well-validated evidence. But **the focus is not on assessing the effects of changing the NAAQS, as should be the case**."
- Dr. Thomas: "I believe the evidence presented in the ISA is suitable for reaching a causal interpretation of the effects of air pollution on human health. **The ISA does not address the implications of potential changes in the NAAQS**; it is my understanding that that will be addressed in the draft Policy Assessment document that I have not seen." (Tony Cox comment: My question was whether the information presented in the ISA is suitable for predicting effects on public health of potential changes in the NAAQS, not whether the ISA or the PA makes such predictions.)

Is each of the causal determinations summarized in Table ES-1 (especially those labeled "causal relationship" or "likely to be causal relationship") the only possible causal determination conclusion that is justified by, or consistent, with current scientific evidence? Could different causal determinations be equally well justified (or better justified) by the information presented, or by the totality of current scientific evidence?

Responses from non-CASAC member consultants:

- Dr. North: "It seems clear that at 2.5 ppm ozone exposure causes respiratory damage in mice. NIOSH scientists have declared 5 ppm "immediately dangerous to life and health" (IDLH) for short-term exposure. But for all the other categories of causality and for ambient ozone exposure levels, **I would like to see projections of what reductions in ambient levels might do to avoid adverse human health impacts**, **and similarly for welfare effects.**"
- Dr. Thomas: "**Any judgment of a "causal" or "likely to be causal" relationship is potentially subject to differences of opinion amongst experts.** It is my opinion that the various determinations summarized in Table ES-1 are well justified by the totality of the evidence, based on what I have read in sections IS.4, the supporting appendices, and my general background knowledge of the field of air pollution epidemiology. I do not claim to have read more than a portion of the ISA or to have an exhaustive knowledge of the substance matter of air pollution epidemiology, however. That said, **I would consider it highly unlikely that any other conclusions would be "equally well justified" or "better justified"** than those reached by the authors of the draft ISA. …

Are there changes in the design, analysis, selection, or interpretation of individual studies or in the ISA's processes for interpreting and summarizing them that would improve the validity, credibility, and transparency of the ISA's scientific reasoning and conclusions?

Responses from non-CASAC member consultants:

- Dr. North: "The overall answer is **yes on all counts**."
- Dr. Thomas: "Obviously **I would welcome wider application of the techniques of causal inference** to observational studies, along the lines of those publications I cited in my response to the draft PM2.5 PA. That said, I believe that the weight of evidence approach used by EPA to evaluate the totality of the evidence, experimental and observational, to be highly appropriate and I have no further suggestions for improvement in that process. … I do not have the comprehensive knowledge of the air pollution epidemiology literature to address whether there are key studies missing, other new evidence, or inadequately considered weaknesses."

Appendix C

Questions for Non-CASAC member Consultants on the Draft Ozone ISA from CASAC Members

Dr. James Boylan

Appendix 1 – Atmospheric Source, Chemistry, Meteorology, Trends, and Background

- Is the discussion on metrics and definitions (Section 1.2) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on sources of U.S. ozone and its precursors (Section 1.3) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on ozone photochemistry (Section 1.4) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on inter-annual variability and longer term trends in meteorological effects on anthropogenic and U.S. background ozone (Section 1.5) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on measurements and modeling (Section 1.6) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on ambient air concentrations and trends (Section 1.7) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on U.S. background ozone concentrations (Section 1.8) accurate and complete? If not, what additional information needs to be included?

Appendix 2 – Exposure to Ambient Ozone

- Is the discussion on exposure concepts (Section 2.2) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on exposure assessment methods (Section 2.3) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on personal exposure (Section 2.4) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on copollutant correlations and potential for confounding (Section 2.5) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on interpreting exposure measurement error for use in epidemiology studies (Section 2.6) accurate and complete? If not, what additional information needs to be included?

Appendix 9 – The Role of Tropospheric Ozone in Climate Effects

- Is the discussion on ozone impacts on radiative forcing (Section 9.2) accurate and complete? If not, what additional information needs to be included?
- Is the discussion on ozone impacts on temperature, precipitation, and climate related variables (Section 9.3) accurate and complete? If not, what additional information needs to be included?

Dr. Tony Cox

Background: *The Ozone ISA states (p. E1) that "Key scientific conclusions (i.e., causality determinations; Section ES.4) are presented and explained. They provide the scientific basis for developing risk and exposure analyses, policy evaluations, and policy decisions for the review. …The ISA thus provides the policy-relevant scientific information that supports the review of the NAAQS."*

Overarching Questions: I have the following overarching methodological questions about the ISA. (I do *not* request direct responses to these overarching questions, although I welcome answers from experts who care to provide them. I do seek answers to the more specific questions that follow these overarching ones.)

1. Is the scientific information provided by the ISA *clear*?
    a. Is it clear how the ISA's causal determination conclusions can be tested, and either verified or refuted (or left undecided), by observations?
    b. Do the concepts and terms used to express key scientific conclusions in the ISA, especially the causal determination categories, have clear scientific meanings (e.g., unambiguous operational definitions)?
    c. Is it clear and generally understood and agreed what the key conclusions mean? Specifically, are the causal determination categories used to communicate key conclusions unambiguous and well defined?
    d. Do those who read the ISA have a shared, unambiguous understanding of what its key scientific conclusions (i.e., causality determinations) imply about how or whether changes in ozone air pollution would change public health outcomes?
2. Is the scientific information provided by the ISA *sound*?
    a. Are its conclusions logically implied by the data and analyses on which they are based?
    b. Are its conclusions correctly stated and caveated?
    c. Is it clear than conclusions do not reflect selection bias in the choice of studies relied on? Are its conclusions consistent with other relevant data and studies not included in the ISA?
    d. Is it clear how studies were selected for inclusion in the ISA, and why individual studies were included or excluded?
    e. Are there other studies that are omitted from the Draft ISA that should be included?
    f. Are the studies relied on to draw conclusions themselves sound (i.e., do their conclusion follow from the data and analyses presented, are potential confounders and modeling biases correctly accounted for, and have other criteria for study quality been systematically and correctly applied?)
    g. Is the implementation of the PECOS approach (p. IS-4 and Appendix 10) and the use of quality assurance and peer review (p. IS-5) adequate to assure that relevant studies were selected and that unsound conclusions were detected and avoided?

3. Are the key scientific conclusions provided by the ISA *scientific*?
    a. Do they deal with testable (and potentially falsifiable) facts about the observable world?
    b. Can the information provided be independently verified? If so, how?
4. Is the scientific information provided by the ISA *policy-relevant*?
    a. Does the information presented address how changing ozone NAAQS would change (probable) public health outcomes?
    b. Is uncertainty about the changes in risks caused by changes in exposure appropriately characterized for use in policy making?
    c. Are effects of other factors that modify the health effects of ozone (e.g., co-exposures, co-morbidities, poverty, lagged daily high and low temperatures, etc.) and of confounders characterized, so that the effects on health from changing ozone alone, without changing other factors, are made clear?

These overarching questions motivate the following more specific questions, on which I seek your advice.

1. **Background:** *The ISA states (p. I-6) that "This ISA draws conclusions about the causal nature of relationships between exposure to ozone and health and welfare effects for categories of related effects (e.g., respiratory effects) by integrating recent evidence across scientific disciplines and building on the evidence from previous assessments. Determinations are made about causation, not just association, and are based on judgments of consistency, coherence, and biological plausibility of observed effects, as well as related uncertainties."* I think "observed effects" here probably means "observed associations," since associations can be calculated from data, but causal effects of exposures (e.g., the difference between the outcome that occurred and the outcome that would have occurred had exposures been different) are inherently unobservable. (Also, consistency, coherence, biological plausibility, and other Bradford Hill considerations apply specifically to associations.) However, associations are not effects (Petitti DB. Associations are not effects. Am J Epidemiol. 1991 Jan 15; 133(2):101-2.) As noted in a recent review, "The field of environmental health has been dominated by modeling associations, especially by regressing an observed outcome on a linear or nonlinear function of observed covariates. Readers interested in advances in policies for improving environmental health are, however, expecting to be informed about health effects resulting from, or more explicitly caused by, environmental exposures. The quantification of health impacts resulting from the removal of environmental exposures involves causal statements. Therefore, when possible, causal inference frameworks should be considered for analyzing the effects of environmental exposures on health outcomes." (Bind MA. Causal Modeling in Environmental Health. Annu Rev Public Health. 2019 Apr 1;40:23-43. doi: 10.1146/annurev-publhealth-040218-044048.)
   Question: My question is: *Can valid determinations of manipulative or interventional causation – that is, how and whether changing exposure would change health risks – be made based on observed associations of the types analyzed in the ISA?* I emphasize manipulative and interventional causation

(rather than predictive (Granger) causation, but-for causation, epidemiological (attributive) causation, mechanistic causation, etc.) because it is most relevant for policy makers. (Further background for this question and references to relevant technical literature are in Cox LA Modernizing the Bradford Hill criteria for assessing causal relationships in observational data. Crit Rev Toxicol. 2018 Nov 15:1-31. doi: 10.1080/10408444.2018.1518404.)

2. **Background:** *The ISA states that "The ISA uses a formal causal framework to classify the weight of evidence using a five-level hierarchy [i.e., 'causal relationship'; 'likely to be a causal relationship'; 'suggestive of, but not sufficient to infer, a causal relationship'; 'inadequate to infer a causal relationship'; 'not likely to be a causal relationship' as described in Table II of the Preamble (U.S. EPA, 2015)] that is based largely on the aspects for causality proposed by Sir Austin Bradford Hill, as well as other frameworks to assess causality developed by other organizations."* I worry that the causal determination categories, and conclusions expressed using them, lack clear scientific meanings. (I understand that they may have clear implications for action, e.g., that effects labeled "causal" or "likely to be causal" may be considered further as grounds for possible revision of NAAQS regulations.) I would like to learn whether expert readers of the ISA have a common, unambiguous understanding of what these categories mean. Since the causal determination framework is used to communicate the key findings of the ISA, I believe it is important to obtain a thorough understanding of what its terms mean.
Questions: The following questions are intended to help assess the conceptual clarity and meaning of the causal determination categories, and of key conclusions expressed using them, such as those in Table ES-1 (p. E-5) of the Draft ISA. Some of the questions are quite specific (of the form "Does this category mean or imply this condition?"). For clarity, *please provide, if possible, an explicit "yes" or "no" answer to each question*, in addition to any further answer you may choose to give. If neither response is appropriate, please provide, if possible, an explicit "unsure of the answer/answer is not clear to me/ unknown" or "answer depends on conditions" in addition to any further answer you may give. Where your answers are "yes" or "no," please provide specific supporting references. Where they are "answer depends on conditions", please specify the conditions on which the answer depends.
    a. A preliminary question: *Is this actually a "formal causal framework*"? (It seems to me to be excessively informal, akin to a Rorschach test, insofar as I have been unable to obtain any formal, unambiguous or operational definitions of the quoted terms 'causal relationship', 'likely to be a causal relationship', 'suggestive of, but not sufficient to infer, a causal relationship', 'inadequate to infer a causal relationship', and 'not likely to be a causal relationship' described in Table II of the Preamble. This seems to me to leave users to make up their own (possibly different) implicit definitions, or to avoid specifying definitions at all. I would greatly appreciate references to any formal, unambiguous operational definitions of these terms. The descriptions in Table II of the Preamble seem to me to be logically incoherent and ambiguous, as discussed further in Cox LA (2019), Improving causal determination. www.sciencedirect.com/science/article/pii/S2590113319300045.)

b. *Does the ISA's causal determination framework clearly distinguish between necessary and sufficient causation?* Does a causal determination of "causal relationship" between ozone exposure and a health response imply that exposure is *sufficient* to cause (or increase the risk of) the response? Does it imply that absence of exposure is sufficient to prevent (or reduce the risk of) the response? If exposure is sufficient to cause a response, but absence of exposure does not reduce the risk of response (because other factors that are always present are also sufficient to cause it), then is the exposure-response relationship still considered "causal" in the ISA framework? (For background, see Gleiss A, Schemper M. Quantifying degrees of necessity and of sufficiency in cause-effect relationships with dichotomous and survival outcomes. Stat Med. 2019 Oct 15;38(23):4733-4748. doi: 10.1002/sim.8331.)
c. Does a determination that exposure has a "causal relationship" with a health effect in a population imply that reducing exposure would reduce risk of the health effect in the population, other factors being held fixed? In other words, *does a "causal relationship" determination imply a manipulative causal relationship?*
d. *Can causal determinations be incorrect*? (Or, to the contrary, are they performative utterances?)
e. If causal determinations can be mistaken, then *is it clear how uncertainty about which category is correct should be (or has been) resolved* in assigning a final causal determination category, as in Table ES-1 p. ES-5) of the ISA?
f. If causal determinations can be incorrect, then *is it clear how observations could be used to test and falsify a given causal determination if it is not correct?* For example, is it completely clear how someone can use relevant data to show that a determination of "causal relationship" or "likely to be causal" in the ISA is incorrect, if indeed that is the case?
g. If causal determinations can be incorrect, then *is the correctness of each causal determination in table ES-1 formally and transparently evaluated in the ISA?* In other words, have formal rules for determining the correctness of the causal determinations in Table ES-1 (p. ES-5) from the data and evidence presented been explicitly stated, applied systematically, and the results documented? (If so, where?)
h. *Does a determination that an exposure-response (or concentration-response (C-R)) relationship is a "causal relationship" imply that it is entirely causal,* with no contribution from incompletely controlled confounding, modeling errors and biases, or other non-causal sources? If not, is there a clearly defined lower bound on how much of the relationship (e.g., how much of the slope of a C-R regression line) must be causal in order for the whole relationship to be classified as causal? (If so, what is it?)
i. *Does a determination that a C-R relationship is a "causal relationship" imply 100% certainty that it is causal?* If not, is there a clearly specified lower bound on how probable it must be that the relationship is causal in order for it to be classified as causal? (If so, what is it?)
j. *Does a determination that a C-R relationship is a "causal relationship" imply that it is causal for every member of a population,* or might it be deemed "causal" if it is causal for a

sensitive subpopulation only? In the latter case, is there a clearly specified lower bound on the fraction of the exposed population for which the relationship must be causal, in order for the whole relationship to be classified as causal? (If so, what is it?)

k. *Are the five categories mutually exclusive?* (Again, please answer yes, no, or unclear, in addition to any other answer you may give. If the answer is yes or no, please cite supporting references.) Might a body of mixed evidence satisfy the definitions for more than one of these categories? (Background for this question is in Cox LA, Improving causal determination. www.sciencedirect.com/science/article/pii/S2590113319300045.) For example, does evidence justifying a causal determination of "causal relationship" preclude (or, conversely, does it imply) a causal determination of "likely to be a causal relationship"? Is evidence categorized as "inadequate to infer a causal relationship" a superset of, a subset of, a disjoint set from, or an overlapping set with, evidence categorized as "suggestive of, but not sufficient to infer, a causal relationship"? Is it possible for a body of evidence to be both "suggestive of, but not sufficient to infer, a causal relationship" and also "inadequate to infer a causal relationship"?

l. *Are the five categories collectively exhaustive?* For example, is evidence satisfying "not a causal relationship" included in any of the five categories? (If so, in which one(s)?) Similarly, is each of the following possible characterizations of evidence compatible with exactly one of the five causal determination categories? (If so, which one? If it is compatible with none, or more than one, then is it clear how one of the five categories should be selected to describe such evidence?)

    i. "likely not to be a causal relationship"
    ii. "likely to be a non-causal relationship" (e.g., a relationship due to confounding or modeling biases)
    iii. "likely to be a predominantly non-causal relationship (e.g., due to residual confounding or to coincident historical trends), but some causal component cannot be ruled out"
    iv. "likely to be a predominantly causal relationship, but some non-causal component cannot be ruled out"
    v. "equally likely to be a causal relationship or a non-causal relationship"
    vi. "more likely than not to be a causal relationship, but evidence is inadequate to infer a causal relationship"
    vii. "likely to have been a causal relationship in the past, when conditions were different, but unlikely to be a causal relationship in the future"
    viii. "likely to be a causal relationship for a few individuals in the population, but not likely to be a causal relationship for the rest of them"
    ix. "causal relationship in the sense of Bradford Hill, but not a causal relationship in the sense of Granger" (more succinctly, "attributive cause but not predictive cause")
    x. "predictive cause but not a mechanistic cause and not a manipulative cause"

m. Can a body of evidence be categorized as "likely to be causal" if the probability of causality based on the evidence is less than 50%?

3. **Background:** The preceding questions essentially ask about whether the scientific information provided by the ISA is *meaningful*, and what the terms used in the ISA to communicate it mean. The following questions (a through g) ask whether the scientific information provided by the ISA is *sound*, i.e., are its conclusions derived by valid inference from true premises? Are the stated conclusions implied by the data and analyses used to support them? Are they consistent with other data and analyses that are at least as good as those selected? Are they appropriately caveated?
    a. *Study selection and interpretation.* **Background:** Appendix 10 of the ISA describes the study selection process for the ISA. To quickly spot check the results, I searched PUBMED for "ozone respiratory effects causal." Three of the top seven articles returned are as follow, shown here with selected conclusions (emphases added):
        - Qian Z, He Q, Lin HM, Kong L, Zhou D, Liang S, Zhu Z, Liao D, Liu W, Bentley CM, Dan J, Wang B, Yang N, Xu S, Gong J, Wei H, Sun H, Qin Z; HEI Health Review Committee. Part 2. Association of daily mortality with ambient air pollution, and effect modification by extremely high temperature in Wuhan, China. Res Rep Health Eff Inst. 2010 Nov;(154):91-217. "Among the gaseous pollutants, we also observed statistically significant associations of mortality with NO, and SO2, and that the estimated effects of these two pollutants were stronger than the PM10 effects. The patterns of NO2 and SO2 associations were similar to those of PM10 in terms of sex, age, and linearity. ***O3 was not associated with mortality***."
        - Cox LA Jr, Popken DA. Has reducing fine particulate matter and ozone caused reduced mortality rates in the United States? Ann Epidemiol. 2015 Mar;25(3):162-73. doi: 10.1016/j.annepidem.2014.11.006. "***There were no significant positive associations between changes in PM2.5 or O3 levels and corresponding changes in disease mortality rates*** between 2000 and 2010, nor for shorter time intervals of 1 to 3 years."
        - Goodman JE, Prueitt RL, Chandalia J, Sax SN. Evaluation of adverse human lung function effects in controlled ozone exposure studies. J Appl Toxicol. 2014 May;34(5):516-24. doi: 10.1002/jat.2905. "Overall, ***these studies do not demonstrate a causal association between ozone concentrations in the range of the current National Ambient Air Quality Standard and adverse effects on lung function***."

    None of these negative results is mentioned in the ISA.

    For studies that that *are* cited in the ISA, I performed the following spot checks and found the following results.

    - Page 3-91 of the ISA states that "A limited number of recent studies provide evidence of an association between long-term exposure to ozone and asthma development in children. … An overview of the evidence is provided below. A recent CHS analysis examined asthma incidence in relation to improved air quality in nine southern California

communities (Garcia et al., 2019). ***Decreases in baseline ozone concentrations in three CHS cohorts, enrolled in 1993, 1996, and 2006, were associated with decreased asthma incidence.***" However, Garcia et al. (2019) actually state that "Among children in Southern California, decreases in ambient nitrogen dioxide and PM2.5 between 1993 and 2014 were significantly associated with lower asthma incidence. ***There were no statistically significant associations for ozone*** or PM10." (Garcia E, Berhane KT, Islam T, McConnell R, Urman R, Chen Z, Gilliland FD. Association of Changes in Air Quality With Incident Asthma in Children in California, 1993-2014. JAMA. 2019 May 21;321(19):1906-1915. doi: 10.1001/jama.2019.5357. Emphasis added.)

- Table 3-3 on "Summary of evidence for a likely to be causal relationship between long-term ozone exposure and respiratory effects" cites the study of Moore et al. (2008) ("Ambient ozone concentrations cause increased hospitalizations for asthma in children: An 18-year study in Southern California") as providing "key evidence" for the ISA's causal determination that there is "a likely to be causal relationship between long-term ozone exposure and respiratory effects." Specifically, Moore et al. is cited as providing "***Consistent evidence of an association between long-term ozone concentrations and hospital admissions*** and ED visits for asthma." Yet, follow-up work by Moore et al. (2012) noted methodological limitations of the 2008 paper (especially, that its results may have resulted from incorrect untested modeling assumptions, rather than from information in the data) and provided and applied an improved methodology ("CMRIER" or "causal models for realistic individualized exposure rules"). A key result was that the previous significant effect of ozone was no longer found. (Moore et al. (2012) state that "The results from the original HRMSM analysis based on the continuous ozone variable estimated with the G-computation method resulted in an estimate of an increase of 1.44e-06 in the proportion of asthma-related hospital discharges for a one-unit increase in ozone. [This is the 2008 study cited in Table 3-3 of the ISA.] ***Unlike results from the HRMSM analysis with the continuous ozone variable, the CMRIER results are not significant.*** Note that the HRMSM analysis was based on G-computation estimation which ***artificially relies on untestable parametric modeling assumptions*** to estimate HRMSM parameters when the ETA assumption is violated. Thus, in this ozone study [the 2008 study cited by the ISA], ***significant results from the G-computation analysis may be a consequence of the approach taken and not solely based on the information in the data***." (Moore KL, Neugebauer R, van der Laan MJ, Tager IB. Causal inference in epidemiological studies with strong confounding. Stat Med. 2012 Jun 15;31(13):1380-404. doi: 10.1002/sim.4469.) This more recent paper is not mentioned in the ISA. The ISA cites the 2008 results as "key evidence" without noting that the authors subsequently revised them in the 2012 paper.

Questions: Based on these spot checks, I have the following questions:

i. *Is it clear that the ISA's study selection process has successfully provided a comprehensive, trustworthy, and unbiased selection of the best available science on ozone and health effects?*

ii. Is it clear why results from Moore (2008) are included and cited as "key evidence" but contrary results from Moore (2012) are excluded? More generally, *is it clear that study inclusion and exclusion criteria were applied systematically and neutrally to identify and select the best and most up-to-date studies to inform the ISA's conclusions?*

iii. *Are there other studies that are omitted from the ISA that should be included?*

iv. *Are there studies included in the ISA that should be omitted* (e.g., because of uncontrolled confounding, obsolete or incorrect modeling assumptions, conclusions dependent on unverified assumptions, ecological fallacy, lack of causally relevant information, lack of design that can support valid causal inferences, or other methodological problems?)

v. *Is it clear that the process followed in selecting and summarizing scientific studies in the ISA was sufficient to assure accurate, unbiased, up-to-date, and trustworthy summaries of the relevant scientific literature to inform causal determination judgments?*

vi. *Do you find in the Executive Summary a clear explanation of the extent to which the key evidence supporting the ISA's causal determinations consists of, is sensitive to, or is derived from unverified modeling assumptions, or from modeling assumptions that more recent literature has found to be incorrect or inadequate? Have you found information in the ISA on sensitivity of causal determination conclusions to untested, uncertain, or incorrect assumptions? (If so, where? See Table Annex 6-1, cf p. 6-67 for a discussion of what should be done. Has it be done, and is it clear what the results were?)*

b. *Were the epidemiological studies used to support the causal determinations summarized in Table ES-1 (p. ES-5) and Figure ES-2 (p. ES-6) appropriately designed and analyzed to provide valid scientific information and valid causal conclusions about effects of possible future interventions (rather than just conclusions about historical statistical associations)?* More specifically, were studies relied on for the "causal" (for short-term respiratory effects) and "likely to be causal" (for short-term and long-term metabolic effects) determinations appropriately designed and analyzed to support valid inferences about manipulative/interventional causality? (See Appendix 3, for a discussion of epidemiological studies. See Table 3-3, p. 3-112, for a "Summary of evidence for a likely to be causal relationship between long-term ozone exposure and respiratory effects.") For these observational studies, *were criteria for valid study design and analysis for causal inference (specifically for interventional causation) explicitly stated, systematically applied, and the results transparently presented?* (If so, where?) For background on such criteria, see Campbell DT, Stanley JC (1963), Experimental and Quasi-Experimental Designs for Research, www.sfu.ca/~palys/Campbell&Stanley-1959-Exptl&QuasiExptlDesignsForResearch.pdf. (My concern here is about whether Table 3-3

and other parts of the ISA seek to draw causal conclusions from non-causal premises and from studies that were neither designed nor analyzed to produce valid causal conclusions or information about effects of future interventions. My key question here is: Is this concern justified?)

c. *Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" adequately controlled for potential confounding and residual confounding by variables such as income and weather variables?* **Background:** (For background on the importance of confounding by temperature, see e.g., Kai et al. (2018), "Does temperature-confounding control influence the modifying effect of air temperature in ozone-mortality associations?" This article concludes that using a categorical variable (e.g., a season indicator) to control for temperature yields highly significant ozone effects at high temperatures, but also significant residual confounding; and that adjusting for (nonlinear) effects of temperatures "substantially reduced ozone effects at high temperatures and residual confounding.") For example, Table 3-3 cites a study by Tétreault et al. as providing "Key Evidence" of "Cohort studies demonstrating an *association* with asthma development in children," which the ISA then interprets as "Evidence for a *likely to be causal* relationship between long-term ozone exposure and respiratory effects."(Emphases added.) In discussing potential confounding, Tétreault et al. state that "We present two confounder models in the results. The first was adjusted for sex and deprivation, whereas the second was adjusted for the same variables as well as the year of birth*."* The article does not mention temperature or weather variables. Tétreault et al. also note their "lack of information on risk factors at the individual level (e.g. socioeconomic status and smoking). We attempted to control for these factors with adjustments of our models using ecological deprivation variables, which are imperfect and ***may result in residual confounding***." (Emphasis added.) **Questions:** *Is the ISA well justified in interpreting the statistical association found by Tétreault et al. as key evidence for a "likely to be a causal relationship"*, given its design and limitations? Is it possible (or plausible) that the association instead reflects uncontrolled or incompletely controlled confounding?

d. *Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" have adequately controlled for biases due to exposure estimation errors or exposure misclassification errors?* For example, Tétreault et al. caution that "First, individual exposure was modeled and not measured through the follow-up, so the quality of the associations depends on the quality of the exposure models. All associations reported in this study were estimated according to the exposure at the centroid of the residential postal code. This assumes that children would stay at home all day. Because a large proportion of a child's day can be spent outside the home (e.g., at school), where exposure to air pollutants might differ, ***misclassification bias may have been introduced in our study***. Additionally, summer average O3 levels were used to estimate annual averages. Because summer O3 levels are higher than winter levels (Environment Canada 1999) in Canada, we may have overestimated annual average levels. Furthermore, although postal

codes circumscribe a relatively small area in urban regions, postal codes may include much larger areas in rural regions. This difference in postal code size could lead to a degree of higher imprecision in exposure estimation in regions of the province that are less densely populated." (Emphasis added.) Does the ISA make adequately clear that the exposure concentrations that it reports (e.g., "32.1 ppb mean summer ozone concentration, based on 8-h midday avg" in Table 3-3) are in fact "modeled and not measured" values? Does it adjust correctly (e.g., using appropriate errors-in-variables methods) for potential biases due to such errors before interpreting the results as key evidence of a likely causal relationship? (If so, where?)

e. *Do you find in the Executive Summary, or elsewhere in the ISA, a clear explanation of the extent to which the key evidence supporting the ISA's causal determinations is sensitive to uncontrolled or incompletely controlled confounding and/or ecological associations?* Page 3-193 of the ISA states that "Sensitivity analyses with alternate specifications for potential confounding inform the stability of findings and aid in judgments of the strength of inference from results." Is it clear how such sensitivity analyses were applied to individual studies (e.g., in interpreting the Tétreault et al. study as adequate to supply "Key Evidence" of a "likely to be causal" relationship)? Is it clear what the results of these sensitivity analyses were? Does the ISA make clear how such sensitivity analyses were used in informing specific causal determinations, and how sensitive the resulting causal determinations are to incompletely controlled confounding? (If so, where?)

f. *More generally, is it clear how criteria for individual study quality were applied to each study used in making causal determinations, and what the results were?* (See Table Annex 6-1, cf p. 6-67.) *Is it clear how the limitations of each individual study were taken into account in causally interpreting their reported associations and in making causal determinations?*

g. *Does the ISA make clear how its causal determinations would change if evidence from associations caused by confounding, residual confounding, measurement error, or unverified modeling assumptions were excluded?*

4. *Is the biological evidence presented in the ISA to support causal determinations correctly stated, correctly interpreted, relevant for predicting effects of changes in the ozone NAAQS, and up-to-date?* For example, should the role of the NLRP3 inflammasome in ozone-induced lung injury be discussed? (See e.g., Michaudel C, Couturier-Maillard A, Chenuet P, Maillet I, Mura C, Couillin I, Gombault A, Quesniaux VF, Huaux F, Ryffel B. Inflammasome, IL-1 and inflammation in ozone-induced lung injury. J Clin Exp Immunol. 2016 Mar 23;5(1):33-40; Xu M, Wang L, Wang M, Wang H, Zhang H, Chen Y, Wang X, Gong J, Zhang JJ, Adcock IM, Chung KF, Li F. Mitochondrial ROS and NLRP3 inflammasome in acute ozone-induced murine model of airway inflammation and bronchial hyperresponsiveness. Free Radic Res. 2019 Jul;53(7):780-790. doi: 10.1080/10715762.2019.1630735.) Is NLRP3 inflammasome activation relevant for ozone risk assessment and for determining whether changes in currently allowed ambient concentrations of ozone would affect public health?

5. *Does the biological evidence presented in the ISA provide well-validated scientific information suitable for predicting the effects on public health of changing NAAQS standard for ozone?*
6. *Is each of the causal determinations summarized in Table ES-1 (especially those labeled "causal relationship" or "likely to be causal relationship") the only possible causal determination conclusion that is justified by, or consistent, with current scientific evidence? Could different causal determinations be equally well justified (or better justified) by the information presented, or by the totality of current scientific evidence?*
7. *Are there changes in the design, analysis, selection, or interpretation of individual studies or in the ISA's processes for interpreting and summarizing them that would improve the validity, credibility, and transparency of the ISA's scientific reasoning and conclusions?*

Dr. Mark Frampton

1. Change in causality determination for short-term cardiovascular effects since the 2013 ISA.

Background: Table ES-1 and section ES.4.1 of the Executive Summary, and Appendix 4, cardiovascular (CV) health effects.

The 2019 ozone ISA has downgraded the causality determination for short-term ozone exposure and cardiovascular effects from "likely" (2013 ISA) to "suggestive". This was due in part to new human clinical studies of CV effects that are inconsistent with the few studies available in 2013, but also to persistent weaknesses in the epidemiological evidence, as reviewed in Appendix 4.1.

Question 1: Please comment on the strengths and weaknesses of the epidemiology literature with regard to CV effects of short-term ozone exposure. Are there key studies that are missing? Are the remaining weaknesses, along with the other new evidence, sufficient to justify the change in causality determination?

2. Metabolic effects, new determination of "likely" for both short- and long-term exposure.

Background: Table ES-1 and section ES.4.1 of the Executive Summary, and Appendix 5, Metabolic Effects.

"Metabolic effects" include effects on body weight, appetite, body composition, caloric intake, diabetes, glucose, insulin, lipid metabolism, stress responses, and thyroid function. Note that "metabolic effects" differ from the issue of metabolic abnormalities as risk factors for other responses. For example, obesity may affect the pulmonary responses to short-term ozone exposures; this should not be considered a "metabolic effect".

This new determination is driven largely by animal toxicology studies, mostly in rodents, and a single human clinical study showing evidence of acute responses in circulating stress hormones.

Question 2: Is there sufficient epidemiological evidence of metabolic effects to justify the "likely" determination for both short- and long-term exposures? Are there additional studies that should be considered?

3. Change in causality determination for total mortality since the 2013 ISA.

Background: Table ES-1 and section ES.4.1 of the Executive Summary, and Appendix 6, Health Effects-Mortality.
The 2019 ozone ISA has downgraded the causality determination for short-term ozone exposure and total mortality from "likely" (2013 ISA) to "suggestive". However, Figure 6-1 on page 6-6, summarizing the epidemiologic studies of short-term total mortality, shows remarkably consistent evidence for an effect. The newer studies are consistent with the findings reviewed in the 2013 ISA.

The rationale for the change is summarized on page 6-20 of the current ISA:

"However, the experimental evidence, specifically from controlled human exposure studies, is not consistent with the studies evaluated in the 2013 Ozone ISA. This contributes additional uncertainty for a biologically plausible mechanism by which short-term ozone exposure could lead to cardiovascular mortality. Lastly, most of the recent studies examined associations between short-term ozone exposure and mortality using ozone data prior to the year 2000, with only Di et al. (2017a) focusing on more recent ozone concentrations."

Although the newer human studies are inconsistent for CV effects, the human studies overall are very consistent for respiratory effects, so there is a plausible pathway for respiratory mortality. In addition, the ISA establishes a new causality category of metabolic effects (see above), with a determination of "likely". Metabolic effects and metabolic syndrome are closely linked with increased risk of CV disease, so this provides a plausible pathway.

Question 3: Please comment on the strengths and weaknesses of the epidemiology literature with regard to short-term ozone exposure and total mortality. Are there key studies that are missing? Does the available evidence justify the change in causality determination for total mortality?

Also please note that, for effects with causal or likely causal determination, the EPA has restricted consideration of epidemiological studies to those in North America (see PECOS Tool, section 6.1.1.1, page 6-3). That was the case for this determination. Are there epidemiological studies of mortality outside of North America that should be considered?

Dr. Sabine Lange

Epidemiology Study Questions

The EPA states in the ISA preamble that "Traditionally, statistical significance is used to a larger extent to evaluate the findings of controlled human exposure and animal toxicology studies. Understanding that statistical inferences may result in both false positives and false negatives, consideration is given to both trends in data and reproducibility of results. Thus, in drawing judgments regarding causality, the U.S. EPA emphasizes statistically significant findings from experimental studies, but does not limit its focus or consideration to statistically significant results in epidemiologic studies."

1) It has been established that associations found in an epidemiology study can be due to: causation, bias, chance, and/or confounding. **If the concept of statistical significance is not useful in epidemiology studies, then how do the study authors/EPA rule out that chance has caused the observed association?**

Some short-term epidemiology studies use a method that is termed "case-crossover". These studies assess the pollutant concentration on the day of a health effect, and "control" days are those days when a person did not experience that health effect. My understanding is that the intention of this method is to control for intra-individual confounders. These study designs often use days before and after the health event (often matched to day of the week) as control days.

2) **Am I correct in understanding that the intention of ozone case-crossover studies is to compare the ozone concentrations on a day when a health effect occurred for a person, to the ozone concentrations on a day when that health effect did not occur for that person?**

3) If so, then it would be important that some other factor (not related to ozone) did not prevent the health event from occurring on a control day. These studies often use days before and after the health event as control days, but for mortality studies (such as Di et al., 2017), how can a day after death be used as a control day? It doesn't matter what the ozone concentrations are after a person's death, that person would not be able to respond to that concentration. **How should we interpret case-crossover studies that use control days after the event (particularly mortality) occurred?**

Experimental Study and Dose Concordance Questions

4) **What is the importance of dose-concordance in establishing the biological likelihood of ozone-mediated effects occurring at relevant exposure concentrations in humans?** Particularly in the context of known dose information about ozone: total inhaled dose includes concentration, exposure time, and exercise duration; Hatch et al., (2013) have shown that humans and rats that are exposed to ozone at rest achieve similar alveolar ozone doses, and that

humans exercising at 5-times a resting ventilation rate achieved an ~ 5-times higher alveolar ozone dose; and that ozone concentrations are 2-10 times lower indoors where people spend most of their time.

5) **Is there evidence that the animal models used to assess ozone effects (largely rats, mice, and non-human primates) are more, less, or similarly sensitive to ozone-mediated adverse effects compared to humans, at approximately equal inhaled doses?**

Causality Question

In this ISA I did not find population studies that considered causal pathways when assessing the association between ozone and health endpoints. It has been shown that the type of interaction between variables (e.g. confounding, colliding, mediating) can impact the results of regression analyses if these variables are controlled for in the regression equation.

6) **In the absence of a causality diagram to direct the choice of variables to control in an epidemiological study, how can we judge whether a study has appropriately controlled for confounders, and has not inappropriately controlled for colliders (which can open up pathways between variables that otherwise would not be connected) or mediators (and thereby controlled away the effect)?**

References

Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J.D., Dominici, F., 2017. Association of Short-term Exposure to Air Pollution With Mortality in Older Adults. Jama 318, 2446. https://doi.org/10.1001/jama.2017.17923

Hatch, G.E., Mckee, J., Brown, J., Mcdonnell, W., Seal, E., Soukup, J., Slade, R., Crissman, K., Devlin, R., 2013. Biomarkers of dose and effect of inhaled ozone in resting versus exercising human subjects: Comparison with resting rats. Biomark. Insights 8, 53–67. https://doi.org/10.4137/BMI.S11102

Dr. Corey Masuca

Appendix 1 Atmospheric Source, Chemistry, Meteorology, Trends, and Background Ozone

1.3.1 Precursor Sources

Are there not other chemicals besides CO and CH4 that also are contained in the precursor mix of ozone formation with its rapidly forming and degradation in the atmosphere?

Does the singling out of these two constituents of the ozone "cocktail" significant as push toward climate change/global warming instead of just evaluation ozone formation?

1.3.1.2.1 Global Methane

Again, is a teasing out/focusing on CH4 important in discussing the virtual "cocktail" of chemicals that may be associated with ozone formation/degradation?

1.3.1.2.2 International Emissions of Ozone Precursors

This section focuses on international transport of ozone precursors.

What about local/state/regional transport of ozone precursors?

1.3.1.3.2 Biogenic Volatile Organic Compounds (VOCs)

It has been stated that biogenic VOCs and contributions are greater than anthropogenic sources (i.e., motor vehicles).

Is there greater confidence in using models and remote sensing (both with relative degrees of uncertainty) to estimate biogenic ozone source contributions that vehicle emissions estimates (manufacturing vehicle emission standards and testing), in making this assessment?

1.4 Ozone Photochemistry

With the advent of monitoring for speciated compounds including PAMS and Near-Road Monitoring (NOy), should there be further discussions about the individual chemicals gleaned from the specialized monitoring.

1.5 Inter-Annual Variability and Longer Term Trends in Meteorological Effects on Anthropogenic and US Background (USB) Ozone

While temperature, wind patterns, cloud cover, and precipitation are highlighted as very important variables in ozone formation, does topography play a role (such as in Birmingham where summertime pollutants are trapped in a "mountainous bowl?"

Are there any independent effects on formation due to relative humidity?

Appendix 2 Exposure to Ambient Ozone

2.3 Exposure Assessment Methods

While monitoring, including fixed, ambient monitors and personal and microenvironmental monitors are highlighted, what about remote sensing? Biological sampling in blood or tissue?

2.3.2.1 Spatial Interpolation

While attempting to quantity concentrations at locations and areas between concentration points is included under **2.3.2 Modeling**, many of these exact same methods (i.e., data averaging, IDW, and kriging) are also utilized for **Monitoring** data shortcomings.

2.4.1 Time-Activity Data

Is it possible that ozone exposure through time-activity data may be reduced due to temperature alone, as more people tend to avoid time spent outdoors in the summers during extremely warm/hot/humid, stagnant days which are oftentimes conditions for greater ozone formation?

Miscellaneous Question(s)

Due to exposure to ozone being disproportionate for disparate (i.e., lower income, children), should this be emphasis in this section, in lieu of regression analysis confounding/covariate in epidemiological studies for low(er) SES?

Dr. Steven Packham

Question 1 Background Statement of Fact:

> Evidence from controlled human exposures is sufficient to conclude with certainty that a causal relationship exists between measurable decrements in FEV1 and subjective symptoms in healthy human adults.

Question 1: *When a causal relationship is conclusive to a high degree of scientific certainty as it is in this case, should this take precedence over causal inference when drafting a NAAQS ISA?*

Question 2 Background Statements of Fact:

1. The shape of the ozone induced FEV1 and subjective symptoms dose-response curve is a function of the inhaled hourly dosage rate and the cumulative dose inhaled over several hours immediately prior to the onset of the effect.
2. The mean cumulative dose threshold for ozone induced FEV1 and symptom effects in healthy adult humans exposed 6.6 hours to ozone concentrations from 60 to 87 ppb is estimated to be 1,362 mg. (Schelegle et al. 2009)
3. This is equivalent to inhaling a dose of 2,439 trillion highly reactive oxidizing molecular moieties.
4. Whatever the oxidative challenge of PM air pollution is to the human lung, it pales in significance to that of ozone.
5. The inhaled hourly dosage rate and cumulative dose thresholds appear to be lower for ozone induced FEV1 and symptom responses than those necessary for inducing clinical signs of injurious pulmonary inflammation.
6. Ozone induced FEV1 decrement and subjective symptoms may be species-specific protective and defensive responses and warning signs for human organisms.
7. Ozone exposures have been shown to stimulate peripheral mucus flow into central bronchi thereby enhancing particle transport from peripheral to central airways and mucociliary clearance of inhaled particulate matter. This beneficial dose dependent response to ozone "…is of interest since it characterizes the reaction of a primary defense mechanism essential to the protection of mucosal surfaces of the tracheobronchial tree." (Forster et al. 1987)

Question 2: *Given evidence available from controlled human exposures substantiating causal relationships with a number of physiological responses, including beneficially confounding interactions of ozone on PM clearance, should Sub-section ES4.1 Health Effects in the Draft's Executive Summary, and the entire Integrated Synthesis section of the Draft be rewritten?*

Question 3 Background Information: Figure ES-3 in the Ozone ISA External Review Draft (shown below) is adapted from the 2013 Ozone ISA which was based on eight human studies published between 1988 and 2013. The 2009 study by Schelegle et al. played a decisive role in the 2015 revision of the O3 NAAQS from 75 to 70 ppb (80 FR 65292 Oct 26, 2015).

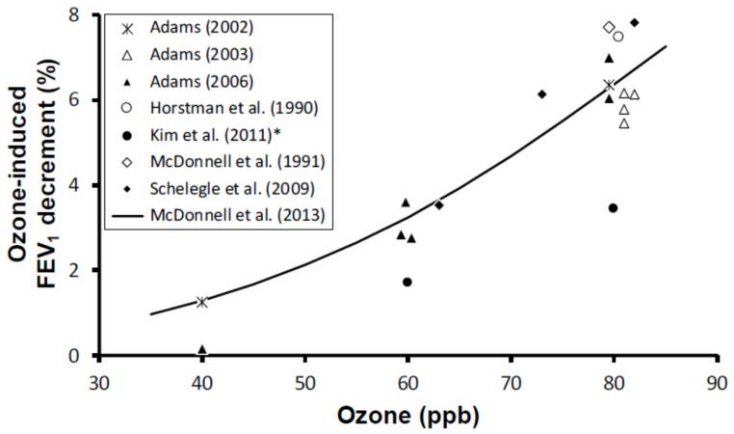


Figure ES-3 was adapted from Figure 6-1 of 2013 Ozone ISA (U.S. EPA, 2013) which was based on studies by Adams (2006), Adams (2003), Adams (2002), Folinsbee et al. (1988), Horstman et al. (1990), Kim et al. (2011), McDonnell et al. (2013), McDonnell et al. (1991), and Schelegle et al. (2009).

Figure 1 below (from Schelegle et al. 2009), on the other hand, depicts the actual mean accumulative doses of 31 healthy adult human subjects who completed four 6.6-hour chamber exposures to target mean O_3 concentrations of 60,70, 80, and 87 ppb. The original data presented *in this way* conveys critical information to toxicologists and biomedical researchers that is "lost in translation" in the concentration/risk-effect picture presented in Figure ES-3.

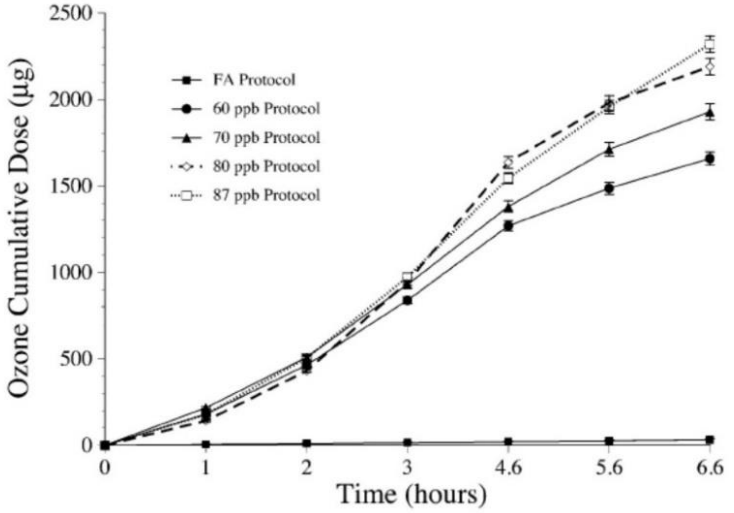
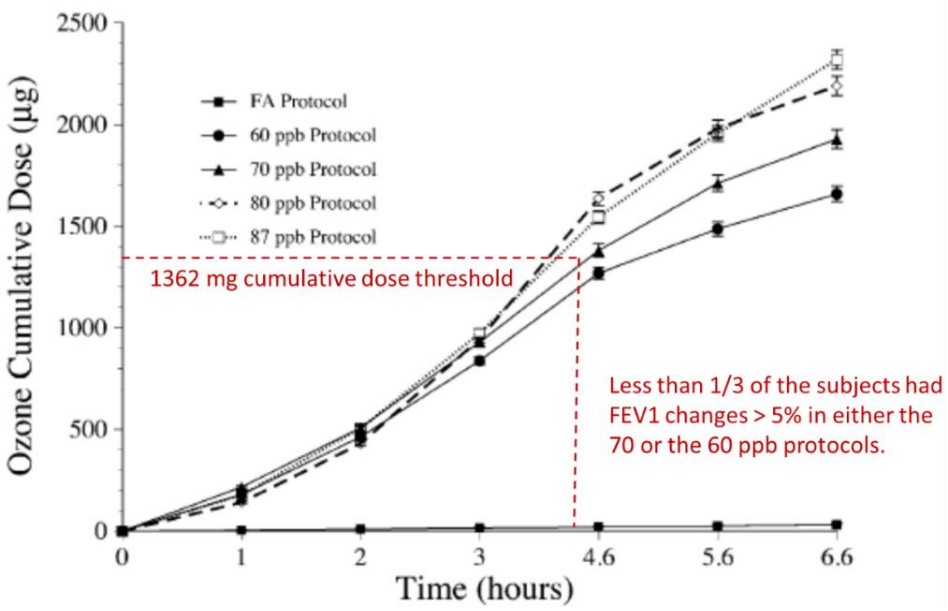


Figure 1. Diagram of mean group values for cumulative dose of ozone (micrograms) against time of exposure for each of the five protocols. Values represent means ± SEM.

To quote Schelegle et al. (2009),

> "We were able to obtain reliable estimates of a Dose of Onset [i.e., a threshold for the FEV1 effect], using the pooled FEV1from the 80 and 87 ppb ozone exposure protocols, …but not from the pooled FEV1 data from the 60 and 70 ppb ozone exposure protocols. The inability to estimate [a threshold] using the FEV1 data from the 60 and 70 ppb ozone exposure protocols is most likely because *less than one third* of the subjects had changes in FEV1 *greater than 5%* in either of these protocols. (Emphasis added)



Packham Figure 1. Adapted from Schelegle et al. (2009) with toxicological annotations by author, 2019. The notable differences between Figure ES-3 compared with Packham Figure 1 are driven by how data are interpreted by different scientific disciplines. By superimposing Schelegle's descriptive conclusion-narrative onto the Sigmoid shaped dose-response curves, one sees the beginning of an increased trend of dose-response curve separation between hour 3 and hour 4: Indicative of the cumulative Dose of Onset threshold between the respective exposure protocols.

Figure ES-3 is the product of adapting (i.e., *imposing*) an ISA Preamble quantal risk-assessment mindset upon graded data collected from continuous response gradients characteristic of living biological organisms. The narrative associated with Figure ES-3 (found on page ES-7) is grossly misleading with respect to the epidemiologically "associated" adverse health effects and completely overlooks the confounding health benefit of enhanced PM clearance stimulated by 200 ppb ozone exposures mentioned above under Question 2 Background Statements of Fact.
The controlled human studies by Folinsbee, Adams, Horstman, Kim, McDonnell and Schelegle, and others cited below in the References and reading list, prove with absolute certainty that exposures to elevated ambient levels of O3 can cause measurable decrements in FEV1 pulmonary test results in healthy adults. These studies document that the effect of O3 on reduced FEV1 volumes is temporary, and suggest that hourly mean ambient O3 concentrations below 70 ppb are not likely to cause FEV1 effects in most healthy adults.

Question 3 Background Statement of Facts:
Several nonmember consultants have expressed reluctance to comment on certain questions because of limited familiarity with pulmonary physiology and inhalation toxicology. Here are few facts to keep in mind.

1. Lungs have an evolutionary history in which surfactant was key to the evolution of all air breathing species on the surface of the planet, (Daniels and Orgeig (2003.)
2. Antioxidant secretions from epithelial Type II cells into the liquid lining of the lungs is one of most important natural defenses the human organism has against naturally occurring ozone levels in the atmosphere near the earth's surface.
3. All known effects of ozone on the human respiratory system are dose dependent.
4. Ozone stimulation of the respiratory airways evokes a number of organism defensive and adaptive responses in humans.
5. Ozone alters tracheobronchial mucociliary function in humans resulting in enhanced transport and clearance of particles deposited in the peripheral air ways, (Foster, et al (1987).
6. Ozone is a potent oxidizing agent, (Pryor et al. (1991).

Question 3 Overarching Conceptual Contexts: An accurate understanding of the causal dose-response relationship between ambient ozone exposure and responses elicited in the human organism opens up a number of important options that could be considered in reviewing and setting NAAQS standards and in how those standards might be used to protect, and even promote, public health. For instance, the realization that the ozone-induced FEV1 effects are temporary, reversible, and occur at a lower inhaled dose than a truly adverse health effect (such as a nonhealing, injurious inflammatory response) could be considered a tenable rationale for classifying them as natural, organism-specific margin-of-safety benchmark indicators.
Another application of hourly MSS inhalation dosage models and thresholds would be to imbed them into web and mobile platform applications for public education and development of user-friendly air quality risk management tools by the EPA. As proof of this concept's possibility, there are two air pollution exposure apps presently in the public domain: A web app http://webapp0.myairhealth.com/# and a free downloadable smartphone app https://apps.apple.com/us/app/myair-health/id790049340.

Question 3: *Looking ahead, do you think toxicology, clinical human studies, and biomedical research disciplines should be given more explicit and balanced consideration in the development of the present, and future, O3 ISAs with the objective to validate causal relationships and determine hourly inhalation dosage rates for adverse inflammatory responses in pulmonary tissues?*

References

Adams WC, Savin WM, Christo AE. (1981). Detection of ozone toxicity during continuous exercise via the effective dose concept. J Appl Physiol 51:415–422.

Adams WC. (2002). Comparison of chamber and face-mask 6.6-hour exposures to ozone on pulmonary function and symptoms responses. Inhal Toxicol 14:745–764.

Adams WC. (2003). Comparison of chamber and face mask 6.6-hour exposure to 0.08 ppm ozone via square-wave and triangular profiles on pulmonary responses. Inhal Toxicol 15:265–281.

Adams WC. (2003). Comparison of chamber and face mask 6.6-hour exposure to 0.08 ppm ozone via square-wave and triangular profiles on pulmonary responses. Inhal Toxicol. 15(3):265-81

Adams WC. (2006a). Comparison of chamber 6.6-h exposures to 0.04-0.08 PPM ozone via square-wave and triangular profiles on pulmonary responses. Inhal Toxicol 18:127–136.

Adams WC. (2006b). Human pulmonary responses with 30-minute time intervals of exercise and rest when exposed for 8 hours to 0.12 ppm ozone via square-wave and acute triangular profiles. Inhal Toxicol 18:413–422.

Bennett WD, Hazucha MJ, Folinsbee LJ, Bromberg PA, Kissling GE, London SJ. (2007). Acute pulmonary function response to ozone in young adults as a function of body mass index. Inhal Toxicol 19:1147–1154.

Bush ML, Asplund PT, Miles KA, Ben-Jebria A, Ultman JS. (1996). Longitudinal distribution of O3 absorption in the lung: gender differences and intersubject variability. J Appl Physiol 81:1651–1657.

Chadha TS, Birch S, Sackner MA. (1987). Oronasal distribution of ventilation during exercise in normal subjects and patients with asthma and rhinitis. Chest 92:1037–1041.

Colucci AV. (1983). Pulmonary dose/effect relationships in ozone exposure. In: Lee SD, Mustafa MG, Mehlman MA, ed. International Symposium on the Biomedical Effects of Ozone and Related Photochemical Oxidants. Princeton, NJ: Princeton Scientific, vol. V, 21–44.

Daniels CB, Orgeig S. (2003). Pulmonary Surfactant: Key to the Evolution of Air Breathing. 0886-1714/03 5.00 © 2003 Int. Union Physiol. Sci./Am. Physiol. Soc. www.nips.org. News Physiol Sci 18:151-157.

Drechsler-Parks DM, Bedi JF, Horvath SM. (1987). Pulmonary function responses of older men and women to ozone exposure. Exp Gerontol 22:91–101.

Folinsbee LJ, Drinkwater BL, Bedi JF, Horvath SM. (1978). The influence of exercise on the pulmonary function changes due to low concentrations of ozone. In: Folinsbee LJ, Wagner J A, Borgia JF Drinkwater BL, Gliner JA, Bedi JF, eds. Environmental Stress. New York: Academic Press, 125–145.

Foster WF, Costa DL, Langenback EG. (1987) Ozone exposure alters tracheobronchial mucociliary function in humans. J Appl Physiol 63:996-1002.

Hazucha MJ. (1987). Relationship between ozone exposure and pulmonary function changes. J Appl Physiol 62:1671–1680.

Hazucha MJ, Folinsbee LJ, Bromberg PA. (2003). Distribution and reproducibility of spirometric response to ozone by gender and age. J Appl Physiol 95:1917–1925.

Horstman DH, Folinsbee LJ, Ives PJ, Abdul-Salaam S, McDonnell WF. (1990). Ozone concentration and pulmonary response relationships for 6.6-hour exposures with five hours of moderate exercise to 0.08, 0.10, and 0.12 ppm. Am Rev Respir Dis 142:1158–1163.

Kim CS, Alexis NE, Rappold AG, Kehrl H, Hazucha MJ, Lay JC, Schmitt MT, Case M, Devlin RB, Peden DB, Diaz-Sanchez D. (2011). Lung function and inflammatory responses in healthy young adults exposed to 0.06 ppm ozone for 6.6 hours. Am J Respir Crit Care Med 183:1215–1221.

Kriebel D, Smith TJ. (1990). A nonlinear pharmacologic model of the acute effects of ozone on the human lungs. Environ Res 51:120–146.

Larsen RI, McDonnell WF, Horstman DH, Folinsbee LJ. (1991). An air quality data analysis system for interrelating effects, standards, and needed source reductions: Part 11. A lognormal model relating human lung function decrease to O3 exposure. J Air Waste Manage Assoc 41:455–459.

McDonnell WF, Stewart PW, Smith MV. (2010). Prediction of ozone induced lung function responses in humans. Inhal Toxicol 22:160–168.

McDonnell WF, Stewart PW, Smith MV. (2007). The temporal dynamics of ozone-induced FEV1 changes in humans: an exposure-response model. Inhal Toxicol 19:483–494.

McDonnell WF, Smith MV. (1994). Description of acute ozone response as a function of exposure rate and total inhaled dose. J Appl Physiol 76:2776–2784.
McDonnell WF, Kehrl HR, Abdul-Salaam S, Ives PJ, Folinsbee LJ, Devlin RB, O'Neil JJ, Horstman DH. (1991). Respiratory response of humans exposed to low levels of ozone for 6.6 hours. Arch Environ Health 46:145–150.

McDonnell WF, Horstman DH, Abdul-Salaam S, Raggio LJ, Green JA. (1987). The respiratory responses of subjects with allergic rhinitis to ozone exposure and their relationship to nonspecific airway reactivity. Toxicol Ind Health 3:507–517.

McDonnell, WF; Stewart, PW; Smith, MV; Kim, CS; Schelegle, ES. (2012). Prediction of lung function response for populations exposed to a wide range of ozone conditions. Inhal Toxicol 24: 619-633. http://dx.doi.org/10.3109/08958378.2012.705919.

McGregor M, Adam W, Sekelj P. (1961). Influence of posture on cardiac output and minute ventilation during exercise. Circ Res 9:1089–1092.

Messineo TD, Adams WC. (1990). Ozone inhalation effects in females varying widely in lung size: comparison with males. J Appl Physiol 69:96–103.

Pryor WA, Das B, Church DF. (1991). The Ozonation of Unsaturated Fatty Acids: Aldehydes and Hydrogen Peroxide as Products and Possible Mediators of Ozone Toxicity. Chem Res Toxicol 4:341-384.

Schelegle, ES; Morales, CA; Walby, WF; Marion, S; Allen, RP. (2009). 6.6-hour inhalation of ozone concentrations from 60 to 87 parts per billion in healthy humans. Am J Respir Crit Care Med 180: 265-272. http://dx.doi.org/10.1164/rccm.200809-1484OC.

Schelegle ES, Adams WC, Walby WF, Marion MS. 2012. Modelling of individual subject ozone exposure response kinetics. Inhal Toxicol 24:401–415.

Seal E Jr, McDonnell WF, House DE, Salaam SA, Dewitt PJ, Butler SO, Green J, Raggio L. (1993). The pulmonary response of white and black adults to six concentrations of ozone. Am Rev Respir Dis 147:804–810.

Silverman F, Folinsbee LJ, Barnard J, Shephard RJ. (1976). Pulmonary function changes in ozone-interaction of concentration and ventilation. J Appl Physiol 41:859–864.

Smith MV, McDonnell WF. (1999). A statistical model for FEV1 response to arbitrary dynamic ozone exposure conditions. J Air Waste Manag Assoc 49:289–298.
Weinmann GG, Weidenbach-Gerbase M, Foster WM, Zacur H, Frank R. (1995). Evidence for ozone-induced small-airway dysfunction: lack of menstrual-cycle and gender effects. Am J Respir Crit Care Med 152:988–996.

Appendix D

Responses to CASAC Member Questions on the Draft Ozone ISA from Non-CASAC member Consultants

Dr. Dan Jaffe, University of Washington-Bothell

Thank you for the opportunity to assist this round of the NAAQS review. Of course I can only respond for questions that are in my area of expertise.

Response to questions from Dr. Sabine Lange:

1) Regarding statistical significance in the preamble, I think the line right above your quoted line in the pre-amble is also important in this context "Other indicators of reliability such as the consistency and coherence of a body of studies as well as other confirming data may be used to justify reliance on the results of a body of epidemiologic studies, even if results in individual studies lack statistical significance." To me, what this is saying that if a group of studies, each with significance at say a 90% confidence all demonstrate a similar effect, then this can be used to justify a conclusion, even if no single study reaches a 95% confidence.

The remaining questions are outside my area of expertise.

Response to questions from Dr. Mark Frampton:

These questions are outside my area of expertise.

Response to questions from Dr. Tony Cox:

These questions are outside my area of expertise.

Response to questions from Dr. Steven Packham:

These questions are outside my area of expertise.

Response to questions from Dr. James Boylan:

I have reviewed Appendix 1 and 2. Overall, I think the ISA is well done, but I have a lot of specific comments and questions. <u>See comments below:</u>

Appendix 1 Atmospheric Source, Chemistry, Meteorology, trends, and background Ozone

Overall comments

- In general, Appendix 1 is well written and accurately conveys the state of our understanding of urban and background O_3.

- The first page bullets, however, are confusing and do not represent a good summary of the section. I recommend these get completely rewritten.
- Some items that need further discussion: (1) impact of background O_3 on high elevation cities in the western US (eg Denver, Salt Lake, etc) (2) the exceptional event rule that allows a state to exclude data when it is deemed out of their control, (3) fact that nationally the declining trend in O_3 since 2013 has greatly slowed (eg Figures 1-9 and 1-10),

Detailed comments

Page 1: This seems like an odd prioritization of key points.

Bullet 1: This really overstates the uncertainties and fails to point out that we have a reasonable understanding of O_3 both in urban and background contexts.

Bullet 3
- These also impact local O3 production due to importance on temperature and stagnation.

Bullet 4
- "U.S. background ozone continues to account for a large fraction"—what does this mean? Too vague and misleading.

Page 1-3

Page 10	Confusing. O3 concentrations are now best expressed as mole fractions (per WMO, see Gaudel 2018) not mixing ratios and these do not depend on temp and pressure.
Line 26	Important to discuss average protocols. Significant figures are important both in the calculation of MDA8s and the ODV.

Page 1-6

Line 32	But this is also true for USB

Page 1-7

Line 1	USB is higher at high elevations. Need discussion of elevation effects on USB and its important for high elevation cities like Denver.
Lines 21-23	Add wildfires here.

Page 1-13

Lines 23-33	Good narrative. Need to discuss approaches, e.g., models informed by observations: E.g. Reidmiller et al., 2009; Zhang et al., 2009. See refs below.
Line 28	"emissions plumes" Confusing phrase.

Page 1-15

Line 25	"high-altitude"Misleading. Usually high altitude refers to >10km.
Line 34	Insert "mountain-top observations," into "plumes have been observed by mountain-top observations, aircraft, sondes,…"

Line 35 Delete "upper"

Page 1-16
Line 2 grammar

Page 1-18
Fig 1-5 From Asia? From China? World?

Page 1-19
Line 11 But still could be important regionally, e.g., CAC, central valley, or SE US

Page 1-21
Line 28 Good summary discussion

Page 1-22
Line 24 Cite original studies: Gong, 2017, and McClure and Jaffe 2018a. See refs below.

Page 1-26
Line 1-2 Make change: "electrical discharge at a voltage sufficient to ~~ionize molecular nitrogen~~ thermally produce nitric oxide (NO)."

Page 1-29
Line 32 grammar

Page 1-31
Line 3 Suggest to cite Jaffe & Zhang 2017 paper on impacts of major high pressure ridge on O3 in California and Pacific Northwest. See refs below.

Page 1-33
Line 22 Make change: "Satellite-based remote sensing methods measure the total ozone column rather than ~~ppm or ppb~~ the in-situ concentration in the atmosphere"
Line 25 Also discuss averaging kernel and vertical sensitivity. Generally, sensitivity in troposphere is limited to one "degree of freedom."

Page 1-34
Line 19 I am not familiar with any analyses that have shown satellites can get surface O3.

Page 1-37
Line 20 For clarity please add info on how "seasons" and "warm season" are defined.

Page 1-40
Table 1-2 title: How is warm season defined?

Page 1-43
Figure 1-7 What is difference between Year-Round Only data and Both Data Sites?

Page 1-45
Figure 1-9 Very little change since 2013

Page 1-46
Line 3 Need to discuss fact that no change since 2013.
Figure 1-10 No change since 2013.

Page 1-48
Line 8 Re: “early afternoon”. Many sites are late afternoon.
Line 9 How widely true is this statement? Not true in LA.

Page 1-49
Line 21 Need to include discussion of elevation of USB, especially for high-elevation cities.

Page 1-56
Line 4 While this is all true in a general sense for seasonal means, really need to discuss episodic USB events and their influence on ODV and exceptional events policy. These are well known to occur due to stratospheric smoke or international pollution.

Page 1-57
Line 1 Add strat

Page 1-61
Line 3-7 But this excludes wildfire impacts, which are part of USB. Recent analyses suggest wildfires are included and this may be (in part) cause for recent lack of O3 decline at some locations. Add references for 2 papers shown inc fire: McClure & Jaffe (2018b) Laing & Jaffe (2019). See refs below.

Suggested references to add:

Gong, X., Kaulfus, A., Nair, U., Jaffe, D.A., 2017. Quantifying O3 impacts in urban areas due to wildfires using a Generalized Additive Model. Environ Sci Technol 51, 13216-13223.

Jaffe, D.A., Zhang, L., 2017. Meteorological anomalies lead to elevated O-3 in the western U. S. in June 2015. Geophys Res Lett 44, 1990-1997.

Laing, J., Jaffe, D., 2019. Wildfires are causing extreme PM concentrations in the western United States. EM: Air and Waste Management Association's Magazine for Environmental Managers, July 2019.

McClure, C.D., Jaffe, D.A., 2018a. US particulate matter air quality improves except in wildfire-prone areas. P Natl Acad Sci USA 115, 7901-7906.

McClure C.D. and Jaffe D.A. 2018b. Investigation of High Ozone Events due to Wildfire Smoke in an Urban Area. Atmos. Envir. https://doi.org/10.1016/j.atmosenv.2018.09.021,

Reidmiller, D.R., Fiore, A.M., Jaffe, D.A., Bergmann, D., Cuvelier, C., Dentener, F.J., Duncan, B.N., Folberth, G., Gauss, M., Gong, S., Hess, P., Jonson, J.E., Keating, T., Lupu, A., Marmer, E., Park, R., Schultz, M.G., Shindell, D.T., Szopa, S., Vivanco, M.G., Wild, O., Zuber, A., 2009. The influence of foreign vs. North American emissions on surface ozone in the US. Atmos Chem Phys 9, 5027-5042.

Zhang, L., Jacob, D.J., Kopacz, M., Henze, D.K., Singh, K., Jaffe, D.A., 2009. Intercontinental source attribution of ozone pollution at western US sites using an adjoint method. Geophys Res Lett 36, L11810, doi: 10.1029/2009GL037950.

Comments on Appendix 2:

Page 2-1

First bullet — The spatial variability depends on the O_3 metric. For example while spatial variability of 24 hour average O_3 is low, variability is much higher for MDA8.

Page 2-2

Line 2 — oxidation of VOCs with NOx as a catalyst.

Line 7 — Mole fraction is now the accepted term. See my comments on Appendix 1 above.

Lines 26-27 — More correctly, VOCs are oxidized and these then react with NOx to form NOx. While NO2 is photolyzed, CO is not photolyzed.

Page 2-5

Line 12 — Sentence is confusing. Fixed monitors will always provide good estimates of ambient concentrations, regardless of spatial variability. This sentence is confusing two diff concepts.

Lines 20-21 — Need to clarify if 6 ppb is referring to 24 hr avg O_3, MDA8 or something else.

Page 2-8

Line 31 — Depends on O_3 metric considered.

Page 2-12

Line 215 — MODIS does not measure O_3. Not sure which instrument this refers to.

Mr. John J. Jansen, Southern Company (retired)

Given the broad nature of the questions posed by the CASAC members, I read a good deal of the draft Ozone Integrated Science Assessment (ES, IS, Appendices 1, 2, and 10, and portions of Appendices 3 through 7). I only scanned appendix 8 since there were no questions posed. I have familiarity with welfare effects and could address it if desired, but not until after November 8 as I have a prior commitment. As with my comments on the PM PA, I respond to many, but not all, CASAC member questions and also offer some general thoughts and other comments at the end.

Questions from Dr. Masuca

As to **precursor sources** page 1-7 notes that the focus of the chapter is on sources of USB ozone so a focus on CO, CH4, **global CH4, and international emissions** is appropriate. There is summary discussion of all ozone precursors and sources and it appears adequate. It would be useful to see some information on the "trend" in biogenic VOC emissions in order to understand its year to year (and season to season) variability. To my knowledge I have never seen such a trend plot and recognize it would be difficult given methodological changes over time but it is in the NEI and might put other pollutant trends into perspective (see figure 1-3c on page 1-11). On page 1-12, I question the statement that mobile sources are the **primary** driver to NOx declines. EGU and industrial sources have also declined significantly.

I am not sure I understand the question on **biogenic VOCs**. Assuming the question's focus is on the CTM's ability to accurately assign responsibility for ozone between biogenic and mobile source VOC, the uncertainties of accurately estimating the specie, amount, location, and timing of both biogenic and mobile VOC are significant. I would submit that although a lot of work has gone into estimating mobile VOC, the limitations are similar to what is summarized in the final bullet for biogenic VOC on page 1-21.

As to **ozone photochemistry**, to the extent that new data from PAMS and near road monitoring is leading to updated CTMs or better CTM performance (or even using the data in model performance evaluation), I would suggest adding such a discussion here or in the modeling discussion (section 1.6). Since the near road data, in particular, is recent, it may be too soon to expect such work and results.

Topography is certainly important locally and needs to be handled in the meteorological models. Humidity has an effect and is handled in the chemical mechanisms of CTMs.

I was disappointed that there was no discussion in section 1.7 on the shifting nature of ozone peak concentrations. With the successful lowering of peak ozone in the summer, the likelihood of the peaks occurring in spring or fall is increasing (see below for references). These findings have implications for monitoring and SIP modeling demonstrations. It may also have implications for source attribution and control strategies if what's needed to reduce peaks in the spring are different from the summer.

- Blanchard, C. L. and G. M. Hidy, "Ozone response to emission reductions in the southeastern United States," Atmos. Chem. And Phys. 18: 8183–8202 (2018).

- Blanchard, C. L., et.al., "Emission influences on air pollutant concentrations in New York State: I. Ozone," Atmospheric Environment: X 3 (2019) 100033

Remote sensing data is discussed relative to USB ozone in section 1.6.1.2. Such data is also used in hybrid modeling for exposure estimation on page 2-12.
I did not understand the remaining questions.

Questions from Dr. Frampton

The three questions basically ask whether specific causality classifications or change in classifications is justified. While this area is not in my area of expertise, I do have some comments about the framework for conducting systematic reviews and making causal determinations (see also responses to Dr. Cox below).

While I commend EPA staff and do not envy them the task of searching, reviewing, summarizing, and evaluating the literature, I have always been frustrated by what I perceive as a lack of clear criteria and transparency in the descriptions of what leads to a particular causality classification. I can read several descriptions of evidence and am unable to identify what makes one "causal" and another "likely" or "suggestive" and as a result have a difficult time deciding whether I agree or not. Clear criteria are needed for study inclusion/exclusion, study quality, and causality classification. It is also not always clear which evidence is being given more weight than other evidence.

These issues has been commented on in many past NAAQS reviews (most recently by Dr. Julie Goodman of Gradient on the Particulate Matter Policy Assessment Document (https://yosemite.epa.gov/sab/sabproduct.nsf//1D9FD74E638BDBAE852584950014F5B5/$File/Goodman+Comments+on+Draft+PM+Policy+Assessment.pdf). The problems persist in this document.

You ask whether additional studies should be included. While I do not know of any, I hope others will offer suggestions. No system is perfect and the review process fills the gap. **In the past**, there appeared to be a tendency to initially include studies with significant positive studies and exclude significant negative as well as null studies. And, therefore, suggestions for missed studies were made. This is not as apparent in this ISA as evidenced by the many figures such as Figure 4-2 on page 4-11. Hopefully any missing studies will get identified in this review.

Again, in the past, when significant negative and null studies were included there was a tendency for the narrative to critique (dismiss?) them and simply accept the significant positive studies. This suggests a bias in the review of the evidence. Identified missing studies also ran the risk of post-rationalization. EPA needs to work harder to critique all studies, weight them appropriately, and avoid post-rationalization of additional studies identified in the review process. Unfortunately, this bias is continuing. See for example page IS-27:

> "While there is coherence between epidemiologic and experimental evidence of ozone-induced lung function decrements and pulmonary inflammation, respiratory symptoms were not associated with ozone exposure in a limited number of epidemiologic studies. However, these

studies generally relied on parent-reported outcomes that may have resulted in under- or over-reporting of respiratory symptoms."

As to your first question, EPA's conclusion that - the weakness in the epidemiologic evidence remains despite some new studies showing significant positive and null results - rest on the lack of significant positive results on for HF, IHD, and MI, arrhythmia and cardiac arrest, or thromboembolic disease. I cannot comment on how critical having such evidence is but it appears to prevent the overall epidemiologic evidence from strengthening. However, looking across all the figures (like 4-2) with relatively few significant positive effects seems to be consistent with "suggestive."

As to question 2, I fall back to the lack of criteria for judging classification. While I am of the school that quality human and animal experimental studies at relevant exposures need to be weighted over suggestive epidemiological (associational) studies to establish causality, I do not know whether there are sufficient suggestive epidemiological studies here. Also, there are no summary figures like figure 4-2 to help me.

As to question 3, figure 6-2 seems to show weak evidence for cardiovascular and respiratory mortality begging the question what is driving the total mortality. The quote from page 6-20 relies on the controlled human exposure studies. In a quick scan, I could not find the controlled human exposure discussion or a figure on those studies.

Questions from Dr. Lange

Question #1: I was not aware of this. Imposing "significance" as a criterion on one type of study and not others seems wrong. On the other hand, does it explain the use of the phrase "positive associations" in many places (e.g., see page 4-2) rather than "significant positive associations?" I find such phrasing troublesome as it could be implying more rigor than exists and lacks clarity. If the result is null, it's null. Significant results should be shown distinctly, not diluted by "positive" null results, and weighted more heavily.

Question #2: I believe that is correct. I also seem to recall use of the ozone on, say, another Tuesday in a month, assuming the event happened on a Tuesday.

Question #3: I believe your concern is not limited to mortality. I would expect using the day after for a hospital admission is affected by medical treatment and being confined indoors. If this is a fatal flaw, the studies should be excluded. If not, then there is clearly an uncertainty and the study results should be down weighted.

Question #6: Clearly, the issue should be a key criteria used in the selection of and evaluation of studies. As I stated above, EPA needs clear criteria for study inclusion/exclusion, study quality, and causality classification. It is also not always clear which evidence is being given more weight than other evidence.

Questions from Dr. Packham

Question #2: The issue of "beneficial" effects should certainly be included. I cannot judge whether a total re-write is needed but balancing adverse versus beneficial effects is always a challenge (e.g., see EPA documents on the SOx, NOx, PM NAAQS secondary effects for nitrogen deposition). Stepping back, part of the problem is choosing which metrics to assess and deciding if the responses are, or lead to, a serious adverse outcome. In the welfare effects literature, it is common to declare any observed response or change as an adverse effect which is problematic. In addition to the issue of beneficial effects, there is the issue of recovery or reversibility. I did a search for the terms in the ISA and found some references to them in discussing experimental studies, which is appropriate. However, I did not see how it affected weighting nor causality classification. In other words, if a metric was responsive but recovered, how is that evidence weighted and used in terms of causality classification?

Question #3: Yes. Clearly Schelegle played a key role last time and there was much debate among the panel, CASAC, and public commenters. And I suspect figure ES-3 is destined for the same fate. Your analysis is helpful and EPA should be encouraged to consider it. As I stated above, I am a proponent of weighting quality human and animal experimental studies at relevant exposures over suggestive epidemiological studies to establish causality.

Questions from Dr. Boylan

Overall, Appendices 1 and 2 do a reasonable job covering the issues. There are a few areas that should be added/addressed namely a) adding trends in biogenic VOC's, b) correcting the description on drivers for NOx trends, c) adding a discussion of near road and PAMS data, and d) adding a discussion on the broadening of when peak ozone is occurring. See my comments on Dr. Masuca's question for more detail on these. Also, there should be discussion on how the various approaches to estimating exposure affect the health analyses. In addition, consistent with my recommendation regarding quantitative uncertainty analysis, these sections should provide a summary of what the quantitative uncertainties are for each.

I am not familiar enough with the literature on ozone effects on climate to render an opinion on the appendix's accuracy and completeness. However, I wonder how relevant the appendix is in setting an ozone NAAQS. There is precious little discussion of US ozone nor US consequences. Since the ultimate purpose of the NAAQS is to protect US public health and welfare, one cannot do so without a US context. How does the information provided allow the Administrator to set a level of US ozone concertation that is protective of US health and welfare related to climate effects? The only "level" information provided is between pre-industrial levels and today's levels. Further, the document explicitly avoids any discussion of feedbacks from ozone precursors and their effects on radiative forcing (see pages 9-4, 9-5, and 9-10). As to effects on temperature, how does one deal with the fact that temperature affects ozone formation and ozone affects temperature. It is also disappointing that no work has been done on the issue of tropospheric ozone effects on UV-B shielding. Finally, there is the issue of change vs. adverse effect. Similar to my concerns with much of the ecological effects literature (especially critical loads literature), only changes are referred to with no perspective. Is the inference that any change is adverse (like in the critical loads literature)? Without such context how does one set a level that protects?

Questions from Dr. Cox

On the issues of the causality framework and causality classification, below are my thoughts:

1. EPA has updated and refined its approach to causality determinations over the years but it has always been qualitative in nature. Despite suggestions to alter and/or add to its approach (especially to add criteria and clarity), EPA has continued to justify the status quo as consistent with past practices endorsed by CASAC.
2. Clear criteria are needed for study inclusion/exclusion, study quality, and causality classification. It is also not always clear which evidence is being given more weight than other evidence.
3. I am of the school that a) associations are not and cannot be causal and b) quality human and animal experimental studies at relevant exposures need to be weighted over suggestive epidemiological (associational) studies to establish causality. In fact, I might argue that robust associational evidence can, at most, infer causality not demonstrate it. The methods advocated by Dr. Cox, once conducted, may or may not alter this statement. To be clear, I am not saying to ignore the associational evidence, only limit how far one interprets it towards causality.
4. In evaluating the strength of the evidence, EPA looks across scientific disciplines (i.e., including epidemiology, controlled human exposure studies, and animal toxicology). This evaluation can lead to a causal classification assuming quality experimental studies are weighted over suggestive epidemiological studies. When EPA evaluates studies across statistical disciplines (i.e., panel, case-crossover, time-series, case-control, and cross-sectional studies), in my opinion, it cannot lead to a causal classification but it can lead to a strengthening of the suggestive associational evidence.
5. It is not always clear which evidence is being given more weight than other evidence. I can read several descriptions of evidence and am unable to identify what makes one "causal" and another "likely" or "suggestive" and as a result have a difficult time deciding whether I agree or not.
6. The assessment, especially of the experimental studies, must deal with beneficial effects, reversibility, and recovery. Stepping back, part of the problem is choosing which metrics to assess and deciding if the responses are, or lead to, a serious adverse outcome. In the welfare effects literature, especially in critical loads research, it is common to declare any observed response or change as an adverse effect which is problematic. In addition to the issue of beneficial effects, there is the issue of recovery or reversibility. I did a search for the terms in the ISA and found some references to them in discussing experimental studies, which is appropriate. However, I did not see how it affected weighting nor causality classification. **In other words, if a metric was responsive but recovered, how is that evidence weighted and used in terms of causality classification?**
7. Finally, I have always been concerned about the term causal vs. the term contribute. We know death certificates are problematic with primary, secondary, and even tertiary causes listed. And I know from personal experience, they are not necessarily correct. How does this factor into the analysis and messaging?

Since I am not a statistician, I cannot respond in detail to your questions. I can only refer to the above. For example, numbers 1 & 2 above would be responsive to your question 2a. As I read through them, some thoughts came to mind. They are listed below:

1. For question 2b, the answer is no simply because the framework does not accommodate those terms. In addition, however, see number 7 above.
2. For question 2h see number 7 above.

3. For question 2l, we need clarity and criteria for the 5 we have, not more. Your list demonstrates the unclear nature and lack of criteria that exists. While it would be nice to have a not causal, I suspect the argument against it is that you cannot prove a negative.
4. For question 3i, your examples show the need for adding studies to the evaluation. No system is perfect and the review process fills the gap as illustrated by your list.
5. For question 3iv, I would suggest that some criteria for minimums on things like confounding would make sense. However, whether a hard line can be established is problematic. Rather, the lack of or minimal confounding analysis should down weight a study.

General Observations

As I stated in my comments on the PM PA I am a proponent of a quantitative uncertainty analysis being performed. And I would argue that the ISA should include a section **in each chapter** on the literature that evaluates the uncertainty in the various components that will make up the risk assessment. One could look to the outline provided in Dr. Lange's question number 4 (questions to consultants on the PM PA) for the beginnings of an outline, adapted of course for ozone. I do believe substantial data and hopefully studies exist to derive estimates for many of the items in her Table 1 and EPA should get on with performing that work. They have been advised to do so in the past. The approaches recommended by Dr. Anne Smith (see below taken from PM PA comments of Dr. North) should also be tried and information embedded in the various studies needed for such analysis should be summarized.

- Smith, A.E. and Gans, W., "Enhancing the Characterization of Epistemic Uncertainties in PM2.5 Risk Analyses, Risk Analysis, 35:361-378 (2015).
- Smith, A.E., "Response to Commentary by Fann et al. on 'Enhancing the Characterization of Epistemic Uncertainties in PM2.5 Risk Analyses,'" Risk Analysis 35:381-384 (2015).
- Smith, A.E., "Inconsistencies in Risk Analyses for Ambient Air Pollutant Regulations," Risk Analysis 36:1737-1744 (2016).
- Smith, A.E., "Response: Author Synthesis and Response," Risk Analysis 36:1780-1792688-1692 (2016).
- Smith, A.E., "Using Uncertainty Analysis to Improve Consistency in Regulatory Assessments of Criteria Pollutant Standards," Perspective, accepted for publication in Risk Analysis (2019). (To become available as soon as typesetting and proofing are done on Early View, https://onlinelibrary.wiley.com/toc/15396924/0/0).

Other Comments

Figure ES-2 on page ES-6 should change the "*" to an up or down arrow to show upgraded and downgraded classification.

EPA should use consistent units when describing a body of evidence. For example, it is frustrating to read, for example, the discussion on pages IS-36 and IS-38 with both ppb and ppm included. It is also misleading.

Dr. Frederick Lipfert, Independent Consultant

Introduction

Ozone differs from other pollutants in important ways. I begin my response with a discussion that illustrates some of them, including nonlinearities that affect the development and interpretation of C-R functions and thus presumption of causality. I also provide some C-R estimates from the literature and some epidemiological background. This information is pertinent to many of the CASAC questions and presenting it here avoids duplication. Appendices include abstracts of 2 papers that may be unfamiliar to CASAC and lists of pertinent references.

Background Discussion of Ozone

Ozone (O_3) is a gaseous pollutant not directly emitted from outdoor sources but created in the *troposphere* by reacting with NO_x and volatile organic compounds (VOCs) in the presence of ultraviolet (UV) light. This is why s*tratospheric* ozone protects against harmful UV radiation. Figure 1 shows the complex patterns of these reactions based on photochemical modeling.

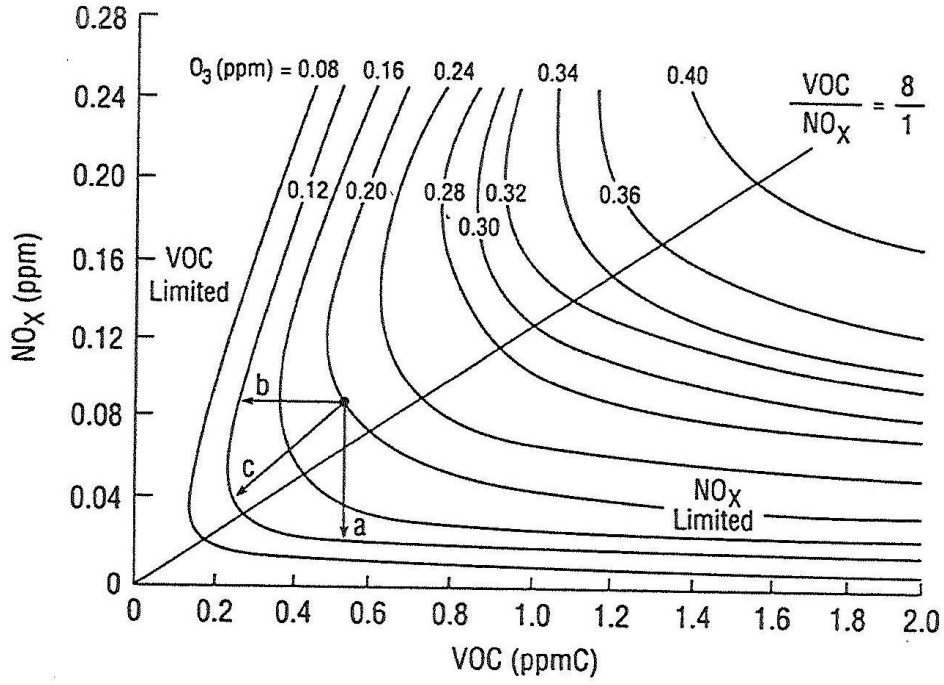


Figure 1. Dependence of ozone on NOx and VOCs (Kleinman and Lipfert, 2003). Current ozone levels (~70 ppb) are near the origin of this plot and the national average NO_x is about 0.02 ppm (Lipfert and Wyzga, 2018). O_3 photochemistry is greatly accelerated at ambient temperatures above 70^0F (Figures 2 and 3), leading to a nonlinear relationship. Above about 80^0F, peak urban O_3 increases by about 4 ppm per degree that would lead to an increase of about 80 ppb at 100^0F relative to about 60 ppb for temperatures around 70^0F. Figure 3 illustrate geographic differences among small cities and rural

areas. Note that O_3 at Mammoth Cave, KY, is about 50 ppb higher on hot days than in North Dakota but that levels tend to be similar under more normal conditions. Time-series studies of acute ozone relationships usually control for temperature, which also has strong nonlinear health effects, especially for the frail elderly. However, only a few long-term studies have attempted to do so, including the Veterans Cohort studies (Lipfert et al., 2000 et seq.)

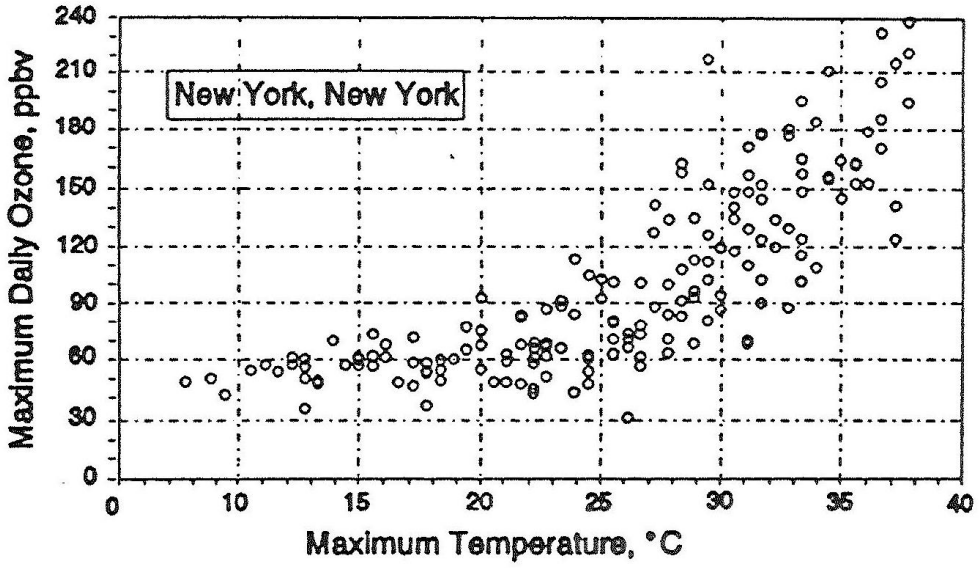


Figure 2. Temperature dependence ozone in New York City (Kleinman and Lipfert, 2003).

Ambient ozone has very different temporal and spatial relationships with respect to other air pollutants, with both diurnal and seasonal peaks (Figure 4). Outdoor workers would thus appear to be most at risk and persons remaining indoors during daylight less so. Urban ozone levels are often high because of high local levels of NO_x and VOCs; those photochemical reactions continue downwind in urban plumes, leading to episodic O_3 levels in suburban and rural areas up to ~100 mi downwind. As a result, exposure patterns based on ambient monitors may be difficult to interpret and gradients in annual. O_3 has no domestic indoor sources and tends to be adsorbed onto indoor surfaces (Kruza et al., 2017) such that indoor levels may be only 20-30% of outdoor. Foley et al. (2003) reported that EPA considered outdoor ambient air quality from fixed ambient monitoring sites to be a "surrogate for exposure".

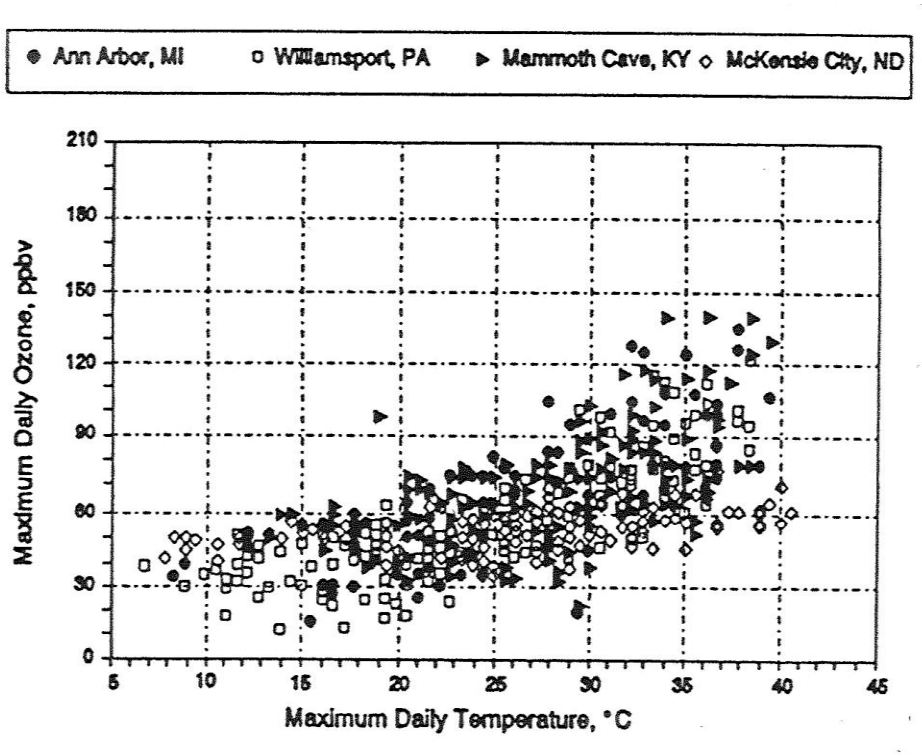


Figure 3. Temperature dependence of daily maximum ozone in 4 U.S. locations

Causality

Epidemiology is concerned with establishing cause and effect, often beginning with statistical associations, but some cautions are in order beyond statistical significance; we need to understand the entire system. The mere presence of a pollutant in the atmosphere does not constitute a "cause", for which personal exposures and translocation to a target organ are required. A long-term "effect" relates to damage to a target organ and implies specific mechanisms; excess mortality or admission to hospital may be a *consequence* of such clinical effects. However, there are exceptions. Acute heat wave mortality cannot always be ascribed to a specific mechanism may result from failure to maintain homeostasis which may also be the case for peak exposures to various air pollutants (Frank and Tankersley, 2002). Cigarette smoking may have different chronic health effects through various mechanisms but has been established as causal. Ammonia and formaldehyde are well-known respiratory irritants.

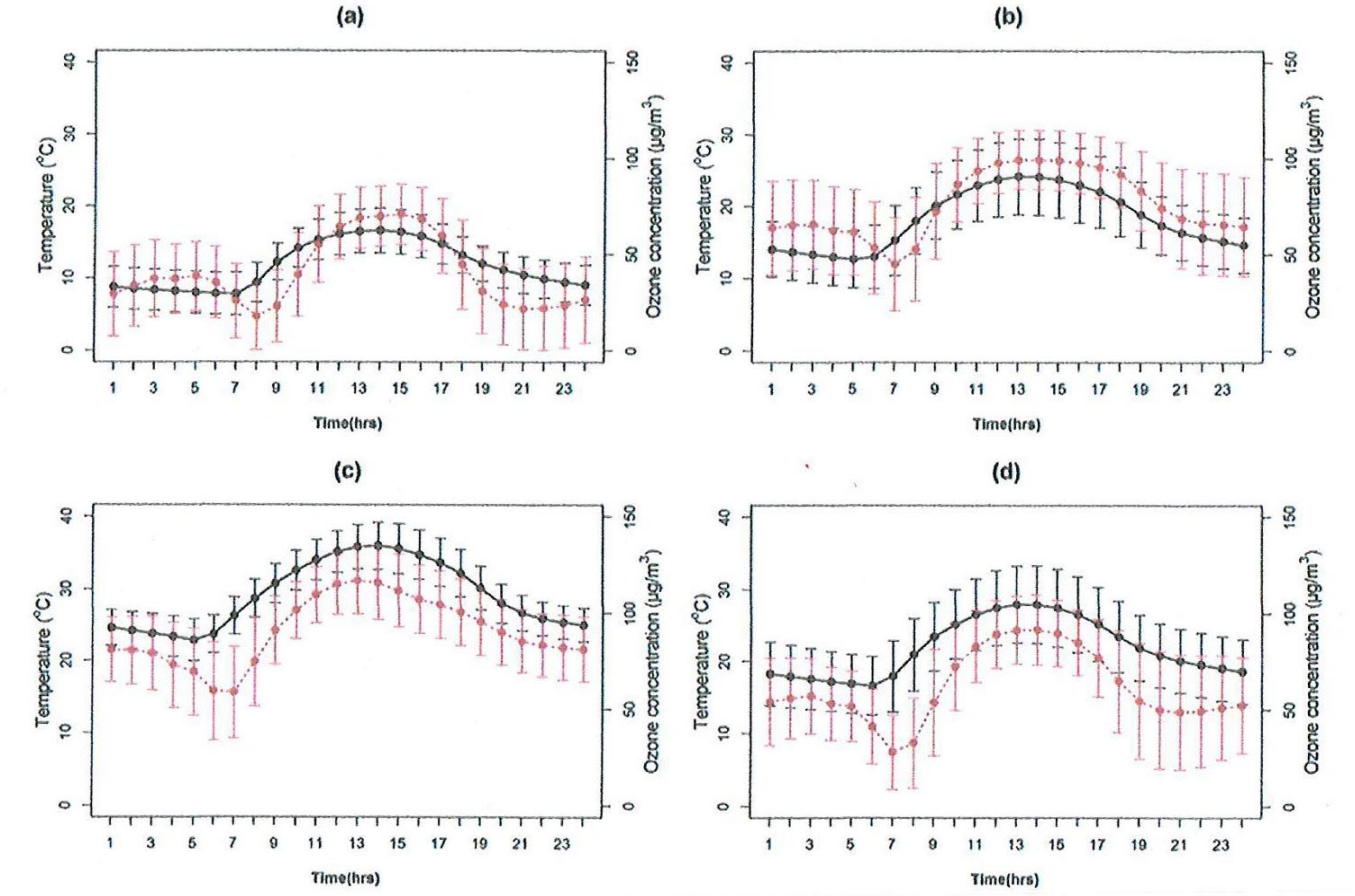


Figure 4. Diurnal variability in temperature (back) and ozone (red) by season. (a) winter. (b) spring. (c) summer (d) autumn. From Pyrgou et al., 2018

Validation of causality through experimental exposures

Comparing health indicators before and after a known exposure is perhaps the most definitive expression of causality and accountability. This was accomplished after the 1952 fog episode in London by autopsies. Time-series analyses in which an endpoint such as daily mortality or admission to hospital rises and falls synchronously with ambient air quality may also qualify after accounting for seasonal cycles, temperature, and co-pollutants. Controlled human exposures may be the most direct means of validation. Figure 5 shows an example for peak ozone.

These experiments were specific to pure ozone without complications arising from ambient co-pollutants such as other photochemical oxidants. They involve precisely determined exposures without measurement error, facilitating identification of thresholds (e.g.,~20 ppb in the above example). Individual test subjects are identified as they would be in a cohort. Lung function is the endpoint in this example and may be related to other endpoints through statistical analysis.

Figure 5 also illustrates potential pitfalls in the assumption of linearity. The data show diminishing benefits from O_3 abatement as the air gets cleaner, as might be expected. However, a linear relationship based on $O_3 > 60$ ppb would have predicted a threshold around 35 ppb, while the nonlinear C-R clearly shows residual risk.

This experimental protocol is problematic for PM, in no small part because ambient PM is not a chemical but rather a mixture that varies in time and place. In previous decades, a device was developed that concentrated ambient particles (CAPS) centrifugally to as much 50 times the local ambient level. Various subjects including humans were tested for various endpoints. However, no pulmonary function impacts were ever found with CAPS experiments (A. Ghio, personal communication Oct. 21, 2019), Graff et al., 2009). It thus appears that health effects on healthy subjects have been validated experimentally for peak ozone but not PM.

However, experiments can provide evidence of health effects under controlled conditions and thus conditions and establish that ozone *may* be harmful to human health. Epidemiology is required to illustrate the conditions under which such health effects *have actually been observed* and the subpopulations most at risk.

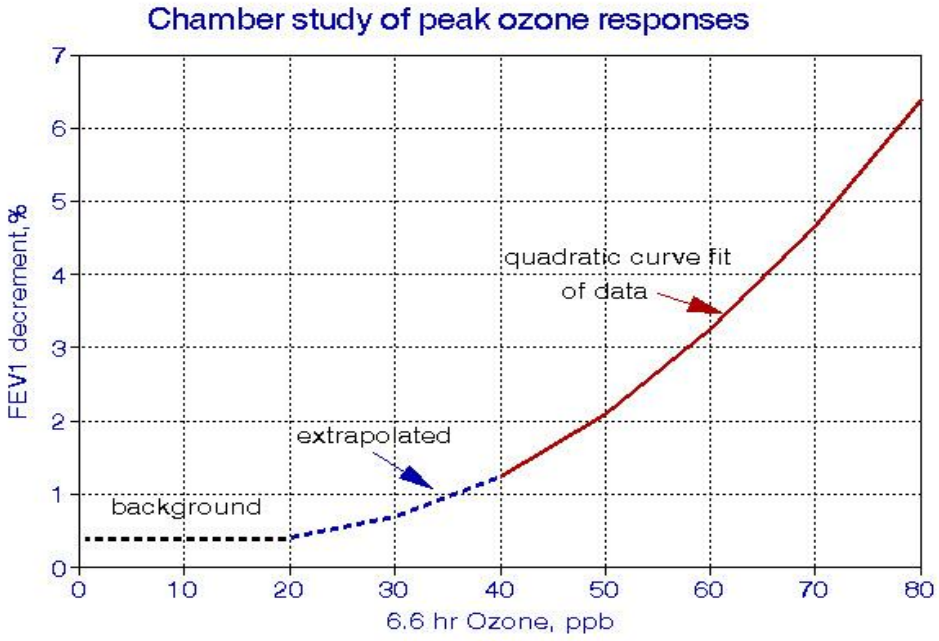


Figure 5. Results of controlled exposures in young healthy adults. Mean decrements in ozone-induced forced expiratory volume in 1 second (FEV_1) after 6.6 hours of exposure. Adapted from Figure ES-3 from the 2019 PA.

Figure 6 compares short-term ozone effects on health subjects' lung function decrements from Figure 5 with daily mortality effects in the U.S. Medicare population (Qian Di et al., 2017). The two C-R relationships are very nearly parallel, implying a 1:1 relationship. However, the FEV_1 data imply a

quasi-threshold (as modified in Figure 1), while the mortality data do not, as predicted by Frank and Tankersley (2002) and shown experimentally in impaired mice. I interpret those experiments as supporting the time-series model of Murray and Lipfert (2012). There is another important distinction between non-lethal experiments on humans and statistical analysis of populations.

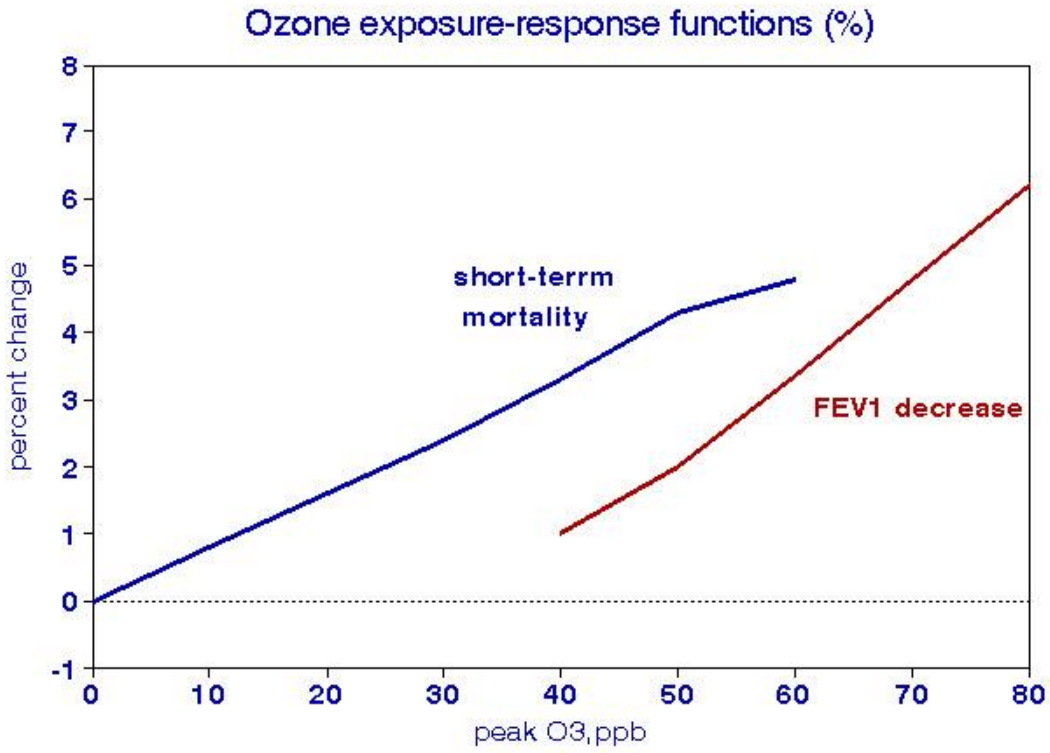


Figure 6. Daily mortality (Qian Di et al., 2017) and lung function decrement (Figure ES-3) as a function of ozone.

Causality in air pollution epidemiology entails five requirements, two of which have been established for peak O_3 (but not long-term mean) levels:

1. Exposure (no). Indoor exposures and daily peaks have not been considered in long-term studies. Initiation of chronic disease only occurs after a latency period and responds to cumulative exposures. This has been shown for smoking but not for air pollution epidemiology. Indoor exposures are not mentioned in Chapter 6.
2. Toxicity (yes). Adverse short-term respiratory effects have been shown in human and animals under specified conditions for selected groups of subjects.
3. Translocation (no). Inhalation of a pollutant is not sufficient to imply contact with an organ and initiation of cardiovascular disease. Associations of O_3 with adverse effects outside the respiratory system are problematic.
4. Susceptibility (yes). Time-series studies have shown that acute mortality is associated with both underlying frailty and daily air pollution peaks. Given the diversity of populations, the notion

that a healthy person could be randomly selected to die coincidently with normal levels of ambient air quality is unreasonable.

5. Accountability (no). The ultimate test of causality is whether health has actually improved since the late 1970s in response to peak O_3 levels reduced by a factor of 5 in conjunction with coincident trends in spatial patterns of reduced smoking and improved medical care. A search of PubMed found no support for such improvement. By contrast Lipfert and Wyzga (2018) showed peak O_3 risk coefficients increasing over time.

Ozone – mortality risk coefficients from the literature

I summarized ozone mortality risk estimates for both long-and short-term exposures from my publications and from the literature in Table 1 (note that my publications have not been included in EPA ISAs. In general, there is a lot of heterogeneity among these estimates, even within the same cohort. However, the long-term estimates from my publications are reasonably self-consistent, more so than those from other studies (Table 1(a)). The overall average long-term risk is 0.076. All of the risk estimates in Table 1 are based on mean O_3 concentrations, not the traditional 10 ppb increment. (Note that health effects of $PM_{2.5}$ have traditionally been based on an increment of 10 μgm^3, which is higher than the national average of 8 10 μgm^3, while the 10 ppb increment for ozone is much lower than the national average.)

The short-term (daily) risk estimates presented in Table 1(b) are based on the average of lags 0 and 1 as has been the custom in many studies. I also show the risk estimates as accumulated up to lag 4, based on Figure 6 (a factor of ~3), with a mean of 0.095. The short-term risk estimates are also quite diverse, leading to the conclusion of no significant difference between long and short terms and thus no truly long-term effects.

I found no evidence of O_3 thresholds in short-term risks (Figure 6) and none would be expected with the Murray frailty model. Since the spatial analyses of long-term risks include short-term risks, the conclusion of no threshold applies to all of the studies. This in turn indicates that the demand of the Clean Air Act for an adequate margin of safety cannot be met.

Table 1. Estimates of ozone-mortality risk coefficients from the literature

(a) Long-term all-cause mortality risks based on the mean value of 95%-ile O_3

1. U.S. ecological study at ages 65-74 (Lipfert and Morris, 2002)

mort period	O_3 period	mean risk	lag (y)
1970-74	1970-74	0.0279	0
1979-81	1979-81	**0.0303**	0
1979-81	1970-74	**0.0415**	7
1989-91	1989-91	**0.0612**	0
1989-91	1979-81	**0.0331**	10
1989-91	1970-74	**0.0405**	17
1995-97	1995-97	**0.158**	0
1995-97	1989-91	**0.0608**	6
1995-97	1979-81	0.0261	16
1995-97	1970-74	**0.0409**	24

2. Veterans Cohort studies (Lipfert et al., 2000)

mort period	O_3 period	mean risk	lag (y)
1976-81	1976-81	**0.102**	0
1982-88	1982-88	**0.146**	0
1982-88	1976-81	**0.100**	7
1989-96	1989-96	0.035	0
1989-96	1982-88	0.060	7
1989-96	1976-81	-0.010	14

Bold = $p < 0.05$, overall mean = **0.0825 (0.66–0.99)**

3. Other long-term studies of all-cause mortality based on mean values of the O_3 metric

location (1st author)	period	O_3 metric	mean risk
all US (Di, 2017)	2000-12	annual mean	0.051
all US (Turner, 2016)	1982-04	8-h max	0.107
all US (Jerrett, 2009)	1977-00	daily max	0.005
California (Jerrett, 2013)	1982-00	monthly av'g	-0.02
Canada (Crouse, 2015)	1991-06	8-h max	0.13

mean risk = 0.055 (-.002-0.112)

Table 1 (b) Short-term mortality risks based on mean values of the O_3 metric

location (1st author)	period O_3 metric		lag 0,1 risk	cumulative risk*
Atlanta (Klemm,	1998-00 max 8-h		0.037-0.064	0.084-0.19
Chicago (Murray, 2012)	1987-00	mean	0.030	0.09
Philadelphia (Lipfert, 2000)	1992-5	mean	0.028-0.04	0.084-0.12
Philadelphia (Lipfert, 2012)	1974-88	daily max	0.014	0.043
US Medicare (Bell, 2004)	1987-00	weekly mean	0.018	0.052
US Medicare (Qian Di, 2018)	2000-12	peak	0.05	0.15
*cumulative risk based on 3x average of lag 0,1			mean risk = **0.095 (0.056-0.13)**	

For comparison, ISA Figure 6-8 shows long-term hazard ratios for 9 North American studies based on 10 ppb that would have a mean risk coefficient of about 0.036 when scaled up a mean concentration of 40 ppb. Three foreign studies had statistically significant negative effects. ISAS Figures 6-1 and 6-2 show short-term mortality risk coefficients for the entire year and for the warm season for all causes, cardiovascular, and respiratory deaths for fixed concentration increments. Warm season risks were substantially higher for respiratory deaths but not total or cardiovascular deaths, in spite of the higher ambient concentrations in summer. which suggests co-pollutant effects. Figure 6-1 shows all-cause all-year risk coefficients for 15 short-term studies that average about 1.5% per standard increment that I estimate as 0.044 (0.012-0.072) on the basis of mean concentrations. The ISA studies are for averages of lags 0-2; estimated cumulative risks would be much larger and similar to the values in Table 2(b) above.

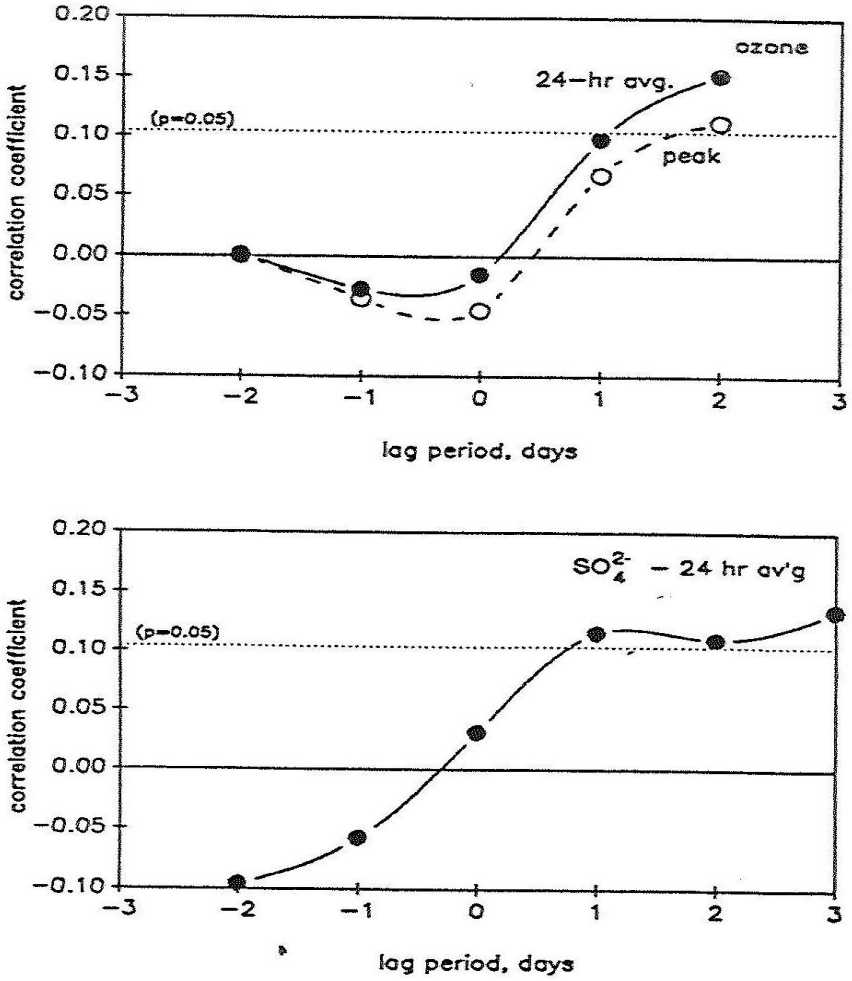


Figure 7. Lag effects on time-series risk coefficients (Lipfert and Hammerstrom, 1992)

Summary of (common sense) ozone epidemiology

I focus on premature mortality as the health endpoint driving cost-benefit analyses. Ozone is a known respiratory irritant associated with symptoms, lung function, and mortality from respiratory diseases. Co-pollutants such as PM may be involved, but PM risks have not been demonstrated experimentally. Ozone exposures vary substantially diurnally and by season but effective control measures must be based on worst-case scenarios and not on long-term averages. Ozone is greatly attenuated indoors but outdoor frequency distributions persist. EPA stated that "exposure has a much more direct link to potential health effects than air quality" (Foley et al., 2003). This principle has been overlooked in more recent years. Indoor exposures are not considered in the discussions of health effects (Chapter 6). Note that peak ozone and temperature are colinear outdoors but not indoors when residential air conditioning is used.

Both short- and long-term associations with all-cause mortality have been shown for specific cohorts and populations, but short-term exposures are included in long-term assessments. Any true long-term effects must thus be based on *differences* between long- and short-term risk estimates. Bona-fide long-term responses that initiate new cases of disease could only could occur as a result of cumulative exposures following a period of latency. Long-term studies in the ISA have not recognized these precepts.

Short-term (daily) effects have been substantially underestimated in the literature. Most were based on the average of the day of death and the following day (lags 0,1), but the true yield of an event must be the sum over the lag period that may be up to three times higher (Figure 7). In addition, indoor concentrations may be only 20-40% of outdoors, such that the risk per unit of personal exposure must be even greater.

Most of the air pollution epidemiology has considered entire cohorts or populations as if all members were at the same risk; however, healthy subjects are at much less risk than those with previous impairments. Short-term deaths are likely to have resulted from loss of homeostasis (Frank and Tankersley, 2002) that may be exacerbated in the frail elderly most at risk. Time-series analyses based on the Murray-Nelson (2000) model have considered frailty as a necessary factor in addition to air pollution and temperature (Murray and Lipfert, 2012). They found that the elderly subpopulations most at risk comprised a small fraction of those aged 65 and over and that their degrees of prematurity were a matter of days, even over exposure periods up to 15 days. These findings have not been considered in ISAs or cost-benefit analyses. The Murray-Nelson model posits that short-term premature death results from the combination of frailty and pollution exposure, such that *either* extreme frailty or high pollution exposure may be responsible. Accordingly, thresholds in either factor alone are proscribed. Linear C-R functions have been shown for short-term ozone exposures as shown in Figure 6.

Table 1 shows that long-term mortality risks estimated from spatial gradients are no larger than those estimated from time-series analyses of short-term peaks. Since the short-term is included in the long term, I conclude that there are no truly long-term mortality effects associated with O_3 exposures, which is consistent with the strong attenuation of O_3 indoors. I also conclude that there is no evidence for a threshold in mortality responses to ozone, which is supported by:

- lack of threshold in lung function decrements in tests of healthy humans
- lack of threshold in daily mortality responses
- importance of frailty in conjunction with ozone in time-series analyses
- confirmation of loss in homeostasis as the driving mechanism in impaired animals.

Figure 5 shows that a NAAQS of 70 ppb implies a lung function decrement of about 4.7%. If we assume a 1:1 relationship between lung function decrement and risk of premature mortality as seen in Figure 6, the mortality risk over background would be 4.3% or 1 in 23, which most people would find excessive. Reducing that risk to 1 in 100 would require a NAAQS of about 42 ppb, which might be very difficult to attain.

My overall conclusion from the above is that a risk-free NAAQS for ozone cannot be determined at any level above background and certainly not as high as 70 ppb. It thus follows that society must determine a

tolerable level of risk, taking into account the presence of extreme frailty. This situation was not considered in the framing of the Clean Air Act.

Summary discussion of the ISA

This is a 1400+ page document for which a thorough review is far beyond the scope of this CASAC question-and-answer investigation. I found the determination of "causality" to be simplistic and notional, as discussed above. As evidence of subjective selection of epidemiology studies, I note that only one of my studies is included (see Table 1 above), but the authors saw fit to mention a study of O_3 exposure and erectile disfunction. (I resisted the opportunity to discuss this further.) Studies with wide confidence intervals were included. There is insufficient linkage between sections, for example between indoor and personal exposures and epidemiology. I found no discussion of accountability, latency in disease incidence, cumulative exposures in epidemiology, comparisons of short- vs. long-term risks, translocation from lung to bloodstream. I conclude that the ISA does not offer convincing evidence of ozone exposure as a cause of long-term health effects in the general population. By contrast, I find the short-term evidence to be more convincing but note that the ISA should compare the two types of studies.

Background references

Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F. Ozone and short-term mortality in 95 US urban communities, 1987-2000. JAMA. 2004 Nov 17;292(19):2372-8.

Crouse DL, Peters PA, Hystad P et al. Ambient PM2.5, O_3, and NO_2 Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ Health Perspect. 2015 ov;123(11):1180-6.

Di Q, Dai L, Wang Y, Zanobetti A, Choirat C, Schwartz JD, Dominici F. Association of Short-term Exposure to Air Pollution With Mortality in Older Adults. JAMA. 2017 Dec 26;318(24):2446-2456.

Di Q, Wang Y, Zanobetti A, Wang Y et al. Air Pollution and Mortality in the Medicare Population. N Engl J Med. 2017 Jun 29;376(26):2513-2522. .

Foley GJ, Georgopoulos PG, Lioy PJ. Accountability within new ozone standards. Environ Sci Technol. 2003 Nov 1;37(21):392A-399A.

Frank R, Tankersley C. Air pollution and daily mortality: a hypothesis concerning the role of impaired homeostasis. Environ Health Perspect. 2002 Jan;110(1):61-5.

Jerrett M, Burnett RT, Pope CA 3rd et al. Long-term ozone exposure and mortality. N Engl J Med. 2009 Mar 12;360(11):1085-95.

Jerrett M, Burnett RT, Beckerman BS et al. Spatial analysis of air pollution and mortality in California. Am J Respir Crit Care Med. 2013 Sep 1;188(5):593-9.

Klemm RJ, Lipfert FW, Wyzga RE, Gust C. Daily mortality and air pollution in Atlanta: two years of data from ARIES. Inhal Toxicol. 2004;16 Suppl 1:131-41.

Kruza M, Lewis AC, Morrison GC, Carslaw N. Impact of surface ozone interactions on indoor air chemistry: A modeling study. Indoor Air. 2017 Sep;27(5):1001-1011.

Lipfert FW, Morris SC, Wyzga RE. Daily mortality in the Philadelphia metropolitan area and size-classified particulate matter. J Air Waste Manag Assoc. 2000 Aug;50(8):1501-13.

Lipfert FW, Perry HM Jr, Miller JP et al. The Washington University-EPRI Veterans' Cohort Mortality Study: preliminary results. Inhal Toxicol. 2000;12 Suppl 4:41-73.

Lipfert FW, Morris SC. Temporal and spatial relations between age specific mortality and ambient air quality in the United States: regression results for counties, 1960-97. Occup Environ Med. 2002 Mar;59(3):156-74.

Lipfert FW, Wyzga RE. Revisiting the Veterans Cohort Mortality Study: New results and synthesis. J Air Waste Manag Assoc. 2018 Nov;68(11):1248-1268.

Lipfert FW. Long-term associations of morbidity with air pollution: A catalog and synthesis. J Air Waste Manag Assoc. 2018 Jan;68(1):12-28.

Murray CJ, Nelson CR. State-space modeling of the relationship between air quality and mortality. J Air Waste Manag Assoc. 2000 Jul;50(7):1075-80.

Murray CJ, Lipfert FW. Revisiting a population-dynamic model of air pollution and daily mortality of the elderly in Philadelphia. J Air Waste Manag Assoc. 2010 May;60(5):611-28.

Murray CJ, Lipfert FW. A new time-series methodology for estimating relationships between elderly frailty, remaining life expectancy, and ambient air quality. Inhal Toxicol. 2012;24(2):89-98.

Murray CJ, Lipfert FW. Inferring frail life expectancies in Chicago from daily fluctuations in elderly mortality. Inhal Toxicol. 2013 Jul;25(8):461-79.

Turner MC, Jerrett M, Pope CA 3rd et al. Long-Term Ozone Exposure and Mortality in a Large Prospective Study. Am J Respir Crit Care Med. 2016 May 15;193(10):1134-42.

Responses to Questions from Dr. James Boylan

Appendix 1 – Atmospheric Source, Chemistry, Meteorology, Trends, and Background

- *Is the discussion on metrics and definitions (Section 1.2) accurate and complete? If not, what additional information needs to be included?*
- *Is the discussion on sources of U.S. ozone and its precursors (Section 1.3) accurate and complete? If not, what additional information needs to be included?*

- *Is the discussion on ozone photochemistry (Section 1.4) accurate and complete? If not, what additional information needs to be included?*
- *Is the discussion on inter-annual variability and longer-term trends in meteorological effects on anthropogenic and U.S. background ozone (Section 1.5) accurate and complete? If not, what additional information needs to be included?*
- *Is the discussion on measurements and modeling (Section 1.6) accurate and complete? If not, what additional information needs to be included?*
- *Is the discussion on ambient air concentrations and trends (Section 1.7) accurate and complete? If not, what additional information needs to be included?*
- *Is the discussion on U.S. background ozone concentrations (Section 1.8) accurate and complete? If not, what additional information needs to be included?*

Sorry, I don't have much to contribute here. I would just point out that for EPA, ambient "ozone" is defined by the "reference" method used to measure it, which applies to inferred health effects as well. It would be of interest to determine how that ambient "ozone" compares to what is used in controlled chamber experiments. It would also be of interest to compare the relationships shown in Figure 1 with actual field experiments.

Background O_3 levels should be discussed in terms of epidemiology and residual risks, including personal and indoor levels, seasonal, and temporal variations, and thresholds. The present discussion of time-activity patterns is inadequate.

Appendix 2 – Exposure to Ambient Ozone

Is the discussion on exposure concepts (Section 2.2) accurate and complete? If not, what additional information needs to be included?

Foley et al. (2003) reported that EPA considered outdoor ambient air quality from fixed ambient monitoring sites to be a "surrogate for exposure". This precept must be directly acknowledged. In this sense, ambient air quality is primarily one of many descriptors of places (county, SMSA, urban vs. rural, etc).

- *Is the discussion on exposure assessment methods (Section 2.3) accurate and complete? If not, what additional information needs to be included?*

- *Is the discussion on personal exposure (Section 2.4) accurate and complete? If not, what additional information needs to be included?*

No. Information on indoor ozone is required including spatial and temporal variations.

- *Is the discussion on copollutant correlations and potential for confounding (Section 2.5) accurate and complete? If not, what additional information needs to be included?*

No. Co-pollutant data (correlations) should include ambient temperature and should be discussed in terms of temporal (hourly) and geographic (urban vs. rural, regional) distributions, especially high vs. low traffic areas.

- *Is the discussion on interpreting exposure measurement error for use in epidemiology studies (Section 2.6) accurate and complete? If not, what additional information needs to be included?*

  The failure to consider indoor ozone is the major source of exposure error. Because of differences in indoor concentrations co-pollutant effects will differ between outdoor and personal exposures.

Appendix 9 – The Role of Tropospheric Ozone in Climate Effects

- *Is the discussion on ozone impacts on radiative forcing (Section 9.2) accurate and complete? If not, what additional information needs to be included?*
- *Is the discussion on ozone impacts on temperature, precipitation, and climate related variables (Section 9.3) accurate and complete? If not, what additional information needs to be included?*

I don't see how this is relevant to setting ambient standards, however interesting.

Responses to Questions from Dr. Tony Cox

Questions on causality: My question is: Can valid determinations of manipulative or interventional causation – that is, how and whether changing exposure would change health risks – be made based on observed associations of the types analyzed in the ISA?

No. See my discussion of causality above.

a. *Is this actually a "formal causal framework"?*

   No, it's a list of subjective rationalizations based on studies selected from the literature according to unspecified procedures. Appendix A lists my own studies not included in the ISA and Appendix B lists the results of relevant PubMed searches.

b. *Does the ISA's causal determination framework clearly distinguish between necessary and sufficient causation?*

   No. In this context there will always be exceptions because of the diversity of the populations at risk. The Clean Air Act anticipated this. This issue involves determination of thresholds.

c. *In other words, does a "causal relationship" determination imply a manipulative causal relationship?*

   Not in my opinion.

d. Can causal determinations be incorrect? (Or, to the contrary, are they performative utterances?)

No, I see causality as binary (yes/no).

e. *If causal determinations can be mistaken, then is it clear how uncertainty about which category is correct should be (or has been) resolved in assigning a final causal determination category, as in Table ES-1 p. ES-5) of the ISA?*

No. I do not.

f. If causal determinations can be incorrect, then is it clear how observations could be used to test and falsify a given causal determination if it is not correct?

Laboratory experiments may determine whether a given type of response *can* happen. Epidemiology is required to determine under what circumstances it *does* happen.

g. If causal determinations can be incorrect, then is the correctness of each causal determination in table ES-1 formally and transparently evaluated in the ISA? In other words, have formal rules for determining the correctness of the causal determinations in Table ES-1 (p. ES-5) from the data and evidence presented been explicitly stated, applied systematically, and the results documented? (If so, where?)

No. I find Table ES-1 useless and completely subjective.

h. Does a determination that an exposure-response (or concentration-response (C-R)) relationship is a "causal relationship" imply that it is <u>entirely</u> causal, with no contribution from incompletely controlled confounding, modeling errors and biases, or other non-causal sources?

I see causality as binary, not conditional.

i. Does a determination that a C-R relationship is a "causal relationship" imply <u>100% certainty</u> that it is causal?

No. My idea of causality demonstration involves physical determination by experiments, any of which may be subject to error and/or misinterpretation. There has been no longitudinal demonstration of public health benefits resulting from the major abatement of ozone exposure since the 1960s.

j. Does a determination that a C-R relationship is a "causal relationship" imply that it is causal for <u>every member</u> of a population, (No) *or might it be deemed "causal" if it is causal for a sensitive subpopulation only?*

Yes, because of the diversity of the population. "Causal" for some portion of the population is still causal but modifiers should be explored

k. Are the five categories mutually exclusive?

Probably not, but they don't merit detailed scrutiny.

l. Can a body of evidence be categorized as "likely to be causal" if the probability of causality based on the evidence is less than 50%?

No, but there are other issues, such as the magnitude of the pooled risk (see below).

8. Study selection and interpretation.

i. Is it clear that the ISA's study selection process has successfully provided a comprehensive, trustworthy, and unbiased selection of the best available science on ozone and health effects?

No. I saw no justification for the studies selected, which included none of mine or my colleagues. (see Appendix A).

ii. Is it clear why results from Moore (2008) are included and cited as "key evidence" but contrary results from Moore (2012) are excluded? More generally, is it clear that study inclusion and exclusion criteria were applied systematically and neutrally to identify and select the best and most up-to-date studies to inform the ISA's conclusions?

No. Moore 2012 cites temperature confounding as a fatal flaw. I would add failure to consider changes that likely varied spatially as well. It seems obvious that temporal changes should be investigated by time-series analysis in a variety of locations. It seems possible that the highly technical aspects of the 2012 Moore paper were beyond the capabilities of the ISA authors (as they are for me as well). These criticisms apply to Tetreault et al. as well.

iii. Are there other studies that are omitted from the ISA that should be included?

Yes. See Appendices A and B to this response.

iv. Are there studies included in the ISA that should be omitted (e.g., because of uncontrolled confounding, obsolete or incorrect modeling assumptions, conclusions dependent on unverified assumptions, ecological fallacy, lack of causally relevant information, lack of design that can support valid causal inferences, or other methodological problems?)

Of course, but one must assume that the most egregious examples would have led to rejection by peer review. My choice would be to retain all relevant studies and to compare their findings by methodological flaw. It is equally inappropriate to summarily reject papers based only on self-selected criteria as it is to accept them in that way.

v. Do you find in the Executive Summary a clear explanation of the extent to which the key evidence supporting the ISA's causal determinations consists of, is sensitive to, or is

derived from unverified modeling assumptions, or from modeling assumptions that more recent literature has found to be incorrect or inadequate? Have you found information in the ISA on sensitivity of causal determination conclusions to untested, uncertain, or incorrect assumptions? (If so, where? See Table Annex 6-1, cf p. 6-67 for a discussion of what should be done. Has it been done, and is it clear what the results were?)

No.

I addressed the following material and questions from Dr. Cox in terms of the ISA's summaries in the executive summary (ES), the integrated summary (IS), and Chapter 6 in the body of the ISA.

b. *Were the epidemiological studies used to support the causal determinations summarized in Table ES-1 (p. ES-5) and Figure ES-2 (p. ES-6) appropriately designed and analyzed to provide valid scientific information and valid causal conclusions about effects of possible future interventions (rather than just conclusions about historical statistical associations)? More specifically, were studies relied on for the "causal" (for short-term respiratory effects) and "likely to be causal" (for short-term and long-term metabolic effects) determinations appropriately designed and analyzed to support valid inferences about manipulative/interventional causality? (See Appendix 3, for a discussion of epidemiological studies. See Table 3-3, p. 3-112, for a "Summary of evidence for a likely to be causal relationship between long-term ozone exposure and respiratory effects.") For these observational studies, were criteria for valid study design and analysis for causal inference (specifically for interventional causation) explicitly stated, systematically applied, and the results transparently presented? (If so, where?) For background on such criteria, see Campbell DT, Stanley JC (1963), Experimental and Quasi-Experimental Designs for Research, www.sfu.ca/~palys/Campbell&Stanley-1959-Exptl&QuasiExptlDesignsForResearch.pdf. (My concern here is about whether Table 3-3 and other parts of the ISA seek to draw causal conclusions from non-causal premises and from studies that were neither designed nor analyzed to produce valid causal conclusions or information about effects of future interventions. My key question here is: Is this concern justified?)*
c. *Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" adequately controlled for potential confounding and residual confounding by variables such as income and weather variables? Background: (For background on the importance of confounding by temperature, see e.g., Kai et al. (2018), "Does temperature-confounding control influence the modifying effect of air temperature in ozone-mortality associations?" This article concludes that using a categorical variable (e.g., a season indicator) to control for temperature yields highly significant ozone effects at high temperatures, but also significant residual confounding; and that adjusting for (nonlinear) effects of temperatures "substantially reduced ozone effects at high temperatures and residual confounding.") For example, Table 3-3 cites a study by Tétreault et al. as providing "Key Evidence" of "Cohort studies demonstrating an association with asthma development in children," which the ISA then interprets as "Evidence for a likely to be causal relationship between long-term ozone exposure and respiratory effects."(Emphases added.) In discussing potential confounding, Tétreault et al. state that "We present two confounder models in the results. The first was adjusted for sex and deprivation, whereas the*

second was adjusted for the same variables as well as the year of birth." The article does not mention temperature or weather variables. Tétreault et al. also note their "lack of information on risk factors at the individual level (e.g. socioeconomic status and smoking). We attempted to control for these factors with adjustments of our models using ecological deprivation variables, which are imperfect and may result in residual confounding." (Emphasis added.)

Questions: *Is the ISA well justified in interpreting the statistical association found by Tétreault et al. as key evidence for a "likely to be a causal relationship", given its design and limitations? Is it possible (or plausible) that the association instead reflects uncontrolled or incompletely controlled confounding?*

d. *Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" have adequately controlled for biases due to exposure estimation errors or exposure misclassification errors? For example, Tétreault et al. caution that "First, individual exposure was modeled and not measured through the follow-up, so the quality of the associations depends on the quality of the exposure models. All associations reported in this study were estimated according to the exposure at the centroid of the residential postal code. This assumes that children would stay at home all day. Because a large proportion of a child's day can be spent outside the home (e.g., at school), where exposure to air pollutants might differ, misclassification bias may have been introduced in our study. Additionally, summer average O3 levels were used to estimate annual averages. Because summer O3 levels are higher than winter levels (Environment Canada 1999) in Canada, we may have overestimated annual average levels. Furthermore, although postal codes circumscribe a relatively small area in urban regions, postal codes may include much larger areas in rural regions. This difference in postal code size could lead to a degree of higher imprecision in exposure estimation in regions of the province that are less densely populated." (Emphasis added.) Does the ISA make adequately clear that the exposure concentrations that it reports (e.g., "32.1 ppb mean summer ozone concentration, based on 8-h midday avg" in Table 3-3) are in fact "modeled and not measured" values? Does it adjust correctly (e.g., using appropriate errors-in-variables methods) for potential biases due to such errors before interpreting the results as key evidence of a likely causal relationship? (If so, where?)*

e. *Do you find in the Executive Summary, or elsewhere in the ISA, a clear explanation of the extent to which the key evidence supporting the ISA's causal determinations is sensitive to uncontrolled or incompletely controlled confounding and/or ecological associations? Page 3-193 of the ISA states that "Sensitivity analyses with alternate specifications for potential confounding inform the stability of findings and aid in judgments of the strength of inference from results." Is it clear how such sensitivity analyses were applied to individual studies (e.g., in interpreting the Tétreault et al. study as adequate to supply "Key Evidence" of a "likely to be causal" relationship)? Is it clear what the results of these sensitivity analyses were? Does the ISA make clear how such sensitivity analyses were used in informing specific causal determinations, and how sensitive the resulting causal determinations are to incompletely controlled confounding? (If so, where?)*

f. *More generally, is it clear how criteria for individual study quality were applied to each study used in making causal determinations, and what the results were? (See Table Annex 6-1, cf p. 6-67.) Is it clear how the limitations of each individual study were taken into account in causally interpreting their reported associations and in making causal determinations?*

g. *Does the ISA make clear how its causal determinations would change if evidence from associations caused by confounding, residual confounding, measurement error, or unverified modeling assumptions were excluded?*

4. *Is the biological evidence presented in the ISA to support causal determinations correctly stated, correctly interpreted, relevant for predicting effects of changes in the ozone NAAQS, and up-to-date? For example, should the role of the NLRP3 inflammasome in ozone-induced lung injury be discussed? (See e.g., Michaudel C, Couturier-Maillard A, Chenuet P, Maillet I, Mura C, Couillin I, Gombault A, Quesniaux VF, Huaux F, Ryffel B. Inflammasome, IL-1 and inflammation in ozone-induced lung injury. J Clin Exp Immunol. 2016 Mar 23;5(1):33-40; Xu M, Wang L, Wang M, Wang H, Zhang H, Chen Y, Wang X, Gong J, Zhang JJ, Adcock IM, Chung KF, Li F. Mitochondrial ROS and NLRP3 inflammasome in acute ozone-induced murine model of airway inflammation and bronchial hyperresponsiveness. Free Radic Res. 2019 Jul;53(7):780-790. doi: 10.1080/10715762.2019.1630735.) Is NLRP3 inflammasome activation relevant for ozone risk assessment and for determining whether changes in currently allowed ambient concentrations of ozone would affect public health?*
5. *Does the biological evidence presented in the ISA provide well-validated scientific information suitable for predicting the effects on public health of changing NAAQS standard for ozone?*
6. *Is each of the causal determinations summarized in Table ES-1 (especially those labeled "causal relationship" or "likely to be causal relationship") the only possible causal determination conclusion that is justified by, or consistent, with current scientific evidence? Could different causal determinations be equally well justified (or better justified) by the information presented, or by the totality of current scientific evidence?*
7. *Are there changes in the design, analysis, selection, or interpretation of individual studies or in the ISA's processes for interpreting and summarizing them that would improve the validity, credibility, and transparency of the ISA's scientific reasoning and conclusions?"*

Comments on the ISA's "Summaries of evidence" and related material relative to Dr. Cox's questions above

These tables (ES-1, for example) comprise the ISA's entire basis for determining causality and therefore the setting of NAAQS levels. I find them subjective and inappropriate for a scientific document. NAAAQS levels affect the entire nation in terms of air pollution control costs and warnings to the public about unhealthy air, and the ISA should use quantitative methods of evaluation for this purpose, rather than subjective judgements.

For example, the ISA used meta-analyses in summarizing welfare effects but not health effects, despite the latter's key role in cost-benefit analysis. The risk estimates in Tables 6-1 and 6-3 and Figures 6-1, 6-2 and 6-8 should be pooled and weighted by inverse variance to derive overall mean estimates and their confidence intervals. This would provide quantitative measures from which the ISA's users could judge for themselves about consistency across studies and hence causality. Consistency (confidence intervals) and any apparent outliers could indicate whether the heterogeneity among studies may be random or the result of study design and the control of confounders. If the latter, subsets of studies having different confounder controls (especially non-linear temperature effects) should be compared.

I note that lags in in Figures 6-1 and 6-2 are mainly 0-2 days while up to 6-d lag is mentioned on p. IS-28. The synthesis (p. IS-1) does not discuss relevant exposure metrics: annual, 8-d, or daily maximum concentrations. Figures 6-1 and 6-2 pooled these metrics using arbitrary increments (15, 20, 25 ppb) without indicating which studies used which increments; these risk estimates are thus ambiguous at best. I recommend using the overall mean concentration to express risk estimates, which is also useful in comparing pollutants, multi-pollutant model results, and short- vs. long-term risks. It is not clear whether any of the short-term estimates summed risk estimates over the lag period, including distributed lags.

The relative magnitudes of the various health effect indicators must be discussed. We would expect them to decrease according to the severity of the effect (larger risks for asthma attacks compared to premature mortality; larger effects on short-term than long-term effects, as discussed above).

Table 6-1. Short-term effects on mortality

The comments below refer to the table's the six sections of decreasing "evidence".

Section 1. It is not clear how "high quality" is determined.

Section 2. Co-pollutant relationships depend on relative measurement errors and say nothing about the validity of either pollutant.

Section 3. "Support for a linear C-R" does not indicate validity. A noisy relationship will appear to be linear in any event. Thresholds can be obscured by indoor air pollution effects.

Section 4. The validity of cardiovascular effects depends on a biological pathway from lung to heart, which has not been shown specifically. Lack of "consistent human exposure studies" is a red flag. In addition, long-term effects cannot be tested experimentally. Figure 6-2 is inconsistent with cardiovascular effects.

Section 5. No comment.

Section 6. Latitude and temperature are important for *all* long-term studies, even the "high quality" ones. Figure 6-4 indicates nonlinear temperature effects and the absence of thresholds which are not adequately discussed

Table 6-2. Long-term effects on mortality

This table comprises four sections of decreasing "evidence" for long-term associations including four subdivisions for the first (Section 1): limited and sometimes "high quality", for which "consistency" seems to be the main criterion. Each of its four subdivisions deals with different models but the overall impression across the four subdivisions is one of inconsistency.

Section 2. Co-pollutant models with varying correlations. Is the implication that the copollutants pose no independent risk? What does this imply for the single-pollutant models? Why are co-pollutants not considered in short-term studies? Co-pollutant effects vary indoors.

Section 3. The implications of linear vs. sublinear C-R functions on causality should be discussed.

Section 4. Biological plausibility should be the primary consideration.

Summary: these comparisons should be quantitative, not qualitative.

Summary comments on causality and study selection.

Cause (causality) is a noun or verb and not an adjective (causal) as used throughout the ISA. It is binary, like pregnancy.

To preclude subjective selections, *all* peer-reviewed studies should be considered regardless of the journal impact factor or the authors' perceived reputations (which I assume has been termed "high quality"); we assume that the peer review process eliminated obvious errors or inappropriate models. (See my catalogue of long-term morbidity studies (Lipfert, 2015); if a single unsupported investigator can sort through ~400 relevant papers, it should not be too much to ask of the ISA authors.) Studies should be grouped by short- vs. long-term, country, endpoint, time period, season, subject characteristics, and (subjective) compliance with the following criteria:

- Is it likely that the exposure estimates used represent personal exposures of the individuals at risk? This requires indoor exposures to be considered.
- Is it likely that all pertinent confounders have been considered, including co-pollutants and climate?
- Is it likely that the pollutants and confounders are linearly related to risk?
- Are there plausible pathways between lung and target organs?
- Are the implied temporal relationships plausible (lags in time-series studies, latency in long-term studies, previous exposures)?
- Are the subjects of the study representative of the general U.S. public?
- Have interactions been properly considered, such as subject age with long-term exposure?
- Does the magnitude of the mean risk pose an important public health concern (1%, no; 10% yes)?

This protocol would shift the judgement process from the authors' overall conclusions to the internal elements of the study. The question thus becomes, given responses to the above criteria, are the authors' conclusions supported, including their estimated confidence intervals? At this point, the ISA could consider to what extent ambient ozone indeed poses health risks and under what conditions, including time scales, based on pooled risk estimates.

The current ozone ISA does not present coherent or convincing evidence for either long- or short-term mortality effects, in large part due to its format and structure.

Questions from Dr. Mark Frampton

1. Change in causality determination for short-term cardiovascular effects since the 2013 ISA.

Background: Table ES-1 and section ES.4.1 of the Executive Summary, and Appendix 4, cardiovascular (CV) health effects. The 2019 ozone ISA has downgraded the causality determination for short-term ozone exposure and cardiovascular effects from "likely" (2013 ISA) to "suggestive". This was due in part to new human clinical studies of CV effects that are inconsistent with the few studies available in 2013, but also to persistent weaknesses in the epidemiological evidence, as reviewed in Appendix 4.1.

Question 1: Please comment on the strengths and weaknesses of the epidemiology literature with regard to CV effects of short-term ozone exposure. Are there key studies that are missing? Are the remaining weaknesses, along with the other new evidence, sufficient to justify the change in causality determination?

The "likely to be causal" categories are lists of subjective rationalizations based on studies selected from the literature according to unspecified procedures; see my general discussion of causality above in which experimental evidence is key. I see mortality causality as binary, ozone can be deadly or not. Morbidity effects should be evaluated in terms of severity and permanence. Epidemiology is then needed to determine under what conditions and how differences among the studies can be explained. In so doing, the time scales of responses are critical. The time-series studies involving frailty by Murray and colleagues have been ignored; Their model posits that frailty is key in short-term mortality by virtue of inability to maintain homeostasis. In that model, differences within the population at risk are key.

2. Metabolic effects, new determination of "likely" for both short- and long-term exposure.

Background: Table ES-1 and section ES.4.1 of the Executive Summary, and Appendix 5, Metabolic Effects.

"Metabolic effects" include effects on body weight, appetite, body composition, caloric intake, diabetes, glucose, insulin, lipid metabolism, stress responses, and thyroid function. Note that "metabolic effects" differ from the issue of metabolic abnormalities as risk factors for other responses. For example, obesity may affect the pulmonary responses to short-term ozone exposures; this should not be considered a "metabolic effect". This new determination is driven largely by animal toxicology studies, mostly in rodents, and a single human clinical study showing evidence of acute responses in circulating stress hormones.

Question 2: Is there sufficient epidemiological evidence of metabolic effects to justify the "likely" determination for both short- and long-term exposures? Are there additional studies that should be considered?

No. The relevance of metabolic effects depends on the demonstration of pathways from the lung to the organs in question. Missing this information, the results are merely associations. Animal and clinical studies are limited to short-term (daily) effects from peak exposures that are unlikely to relate to translocated mechanisms and subsequent longer-term responses.

3. Change in causality determination for total mortality since the 2013 ISA.

Background: Table ES-1 and section ES.4.1 of the Executive Summary, and Appendix 6, Health Effects-Mortality.

The 2019 ozone ISA has downgraded the causality determination for short-term ozone exposure and total mortality from "likely" (2013 ISA) to "suggestive". However, Figure 6-1 on page 6-6, summarizing the epidemiologic studies of short-term total mortality, shows remarkably consistent evidence for an effect. The newer studies are consistent with the findings reviewed in the 2013 ISA. The rationale for the change is summarized on page 6-20 of the current ISA:

"However, the experimental evidence, specifically from controlled human exposure studies, is not consistent with the studies evaluated in the 2013 Ozone ISA. This contributes additional uncertainty for a biologically plausible mechanism by which short-term ozone exposure could lead to cardiovascular mortality. Lastly, most of the recent studies examined associations between short-term ozone exposure and mortality using ozone data prior to the year 2000, with only Di et al. (2017a) focusing on more recent ozone concentrations."

Although the newer human studies are inconsistent for CV effects, the human studies overall are very consistent for respiratory effects, so there is a plausible pathway for respiratory mortality. In addition, the ISA establishes a new causality category of metabolic effects (see above), with a determination of "likely". Metabolic effects and metabolic syndrome are closely linked with increased risk of CV disease, so this provides a plausible pathway.

Question 3: Please comment on the strengths and weaknesses of the epidemiology literature with regard to short-term ozone exposure and total mortality. Are there key studies that are missing? Does the available evidence justify the change in causality determination for total mortality?

Also please note that, for effects with causal or likely causal determination, the EPA has restricted consideration of epidemiological studies to those in North America (see PECOS Tool, section 6.1.1.1, page 6-3). That was the case for this determination. Are there epidemiological studies of mortality outside of North America that should be considered?

I have existential problems with the evaluation of "causality", which has a double exclusion criterion here: a study must report significant positive risks *and* take place in North America. Since Mexico is included, ethnic differences are involved as well as climatic. I would prefer that *any* study deemed to have sufficient subjects and adequate exposure data be considered, in part to specifically evaluate the roles of climate, traffic density, etc. I searched PubMed for citations with "ozone", "daily", and '"mortality" in abstract or title and found 279 papers of which only 47 involved U.S. populations, only 3 of which were published after 2015. See my Appendix B for a list of the 47 short-term papers and 25 long-term. I selected them by abstract and it is not clear how many of them included independent risk

estimates using a valid model. In any event, it is clear that the ISA neglected the overwhelming majority of the available literature of short-term effects on mortality.

Questions from Dr. Sabine Lange

Epidemiology Study Questions

The EPA states in the ISA preamble that "Traditionally, statistical significance is used to a larger extent to evaluate the findings of controlled human exposure and animal toxicology studies. Understanding that statistical inferences may result in both false positives and false negatives, consideration is given to both trends in data and reproducibility of results. Thus, in drawing judgments regarding causality, the U.S. EPA emphasizes statistically significant findings from experimental studies, but does not limit its focus or consideration to statistically significant results in epidemiologic studies."

1) *It has been established that associations found in an epidemiology study can be due to: causation, bias, chance, and/or confounding. If the concept of statistical significance is not useful in epidemiology studies, then how do the study authors/EPA rule out that chance has caused the observed association?*

Essentially, they do not. I take strong issue with the above EPA protocol on statistical significance and note that studies with wide confidence intervals are included. Experimental studies involve defined exposures with no co-pollutants or temperature effects and no measurement error. Variations among studies are thus real and should be analyzed as such. By contrast, epidemiological studies are subject to all of these sources of bias, especially temperature and the treatment of lags. Foley et al. (2003) reported that EPA considered outdoor ambient air quality from fixed ambient monitoring sites to be a "surrogate for exposure", which requires consideration of indoor exposures, by contrast with experimental studies.

Some short-term epidemiology studies use a method that is termed "case-crossover". These studies assess the pollutant concentration on the day of a health effect, and "control" days are those days when a person did not experience that health effect. My understanding is that the intention of this method is to control for intra-individual confounders. These study designs often use days before and after the health event (often matched to day of the week) as control days.

2) *Am I correct in understanding that the intention of ozone case-crossover studies is to compare the ozone concentrations on a day when a health effect occurred for a person, to the ozone concentrations on a day when that health effect did not occur for that person?*

Yes.

3) *If so, then it would be important that some other factor (not related to ozone) did not prevent the health event from occurring on a control day. These studies often use days before and after the health event as control days, but for mortality studies (such as Di et al., 2017), how can a day after death be used as a control day? It doesn't matter what the ozone concentrations are after a person's death, that person would not be able to respond to that concentration. How should we*

interpret case-crossover studies that use control days after the event (particularly mortality occurred?

I am not an expert on case-control studies but I question their use in temporal rather than spatial studies. Ozone has both diurnal and seasonal trends as does ambient temperature, its primary confounder (see the Figures 2 and 3 above). Further, acute effects on mortality persist for several days, perhaps up to a week. Thus "case" days may not be independent of "control" days. Different people with different characteristics die on each day for different reasons. I would rather see both case and control periods extended for say, 3 days or more. The time-series model of Murray and colleagues (not cited in the ISA) considers temporal patterns of subject frailty, ambient temperature and air quality, each over several days. These temporal patterns are much easier to interpret than case-control findings. Di et al. combined spatial and temporal analyses, thus introducing geographic and climate variability (for no particular reason). I would have much rather seen a conventional time-series analysis in each of several locations involving socioeconomic differences as well.

Experimental Study and Dose Concordance Questions

Particularly in the context of known dose information about ozone: total inhaled dose includes concentration, exposure time, and exercise duration; Hatch et al., (2013) have shown that humans and rats that are exposed to ozone at rest achieve similar alveolar ozone doses, and that humans exercising at 5-times a resting ventilation rate achieved an ~ 5-times higher alveolar ozone dose; and that ozone concentrations are 2-10 times lower indoors where people spend most of their time.

4) *What is the importance of dose-concordance in establishing the biological likelihood of ozone-mediated effects occurring at relevant exposure concentrations in humans?*

I fail to see any relevance. Controlled animal and human clinical studies serve only to show what *might* happen under controlled and idealized conditions. Null findings may thus be the most important. By contrast, epidemiology shows what actually *does* happen in the real world, including variability.

5) *Is there evidence that the animal models used to assess ozone effects (largely rats, mice, and non-human primates) are more, less, or similarly sensitive to ozone-mediated adverse effects compared to humans, at approximately equal inhaled doses?*

I don't know but, in my view, such tests should be only qualitative and used to study mechanisms. I don't see them useful to look for a "safe" dose since the actual human doses used in epidemiology remain unknown in part because of indoor effects.

Causality Question

In this ISA I did not find population studies that considered causal pathways when assessing the association between ozone and health endpoints. It has been shown that the type of interaction between variables (e.g. confounding, colliding, mediating) can impact the results of regression analyses if these variables are controlled for in the regression equation.

6) *In the absence of a causality diagram to direct the choice of variables to control in an epidemiological study, how can we judge whether a study has appropriately controlled for confounders, and has not inappropriately controlled for colliders (which can open up pathways between variables that otherwise would not be connected) or mediators (and thereby controlled away the effect)?*

I see causality and confounders in simpler terms: causality in terms of experiments and physiology and confounders in terms of bi-variate correlations and model evaluations with and without potential confounders. As stated above I believe that most of the mortality relationships are short-term and thus with few potential confounders. I don't think one can ever be sure that all of the long-term potential (spatial) confounders have been adequately controlled.

References

Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J.D., Dominici, F., 2017. Association of Short-term Exposure to Air Pollution With Mortality in Older Adults. Jama 318, 2446. https://doi.org/10.1001/jama.2017.17923

Hatch, G.E., Mckee, J., Brown, J., Mcdonnell, W., Seal, E., Soukup, J., Slade, R., Crissman, K., Devlin, R., 2013. Biomarkers of dose and effect of inhaled ozone in resting versus exercising human subjects: Comparison with resting rats. Biomark. Insights 8, 53–67. https://doi.org/10.4137/BMI.S11102

Questions from Dr. Corey Masuca

Appendix 1 Atmospheric Source, Chemistry, Meteorology, Trends, and Background Ozone

1.3.1 Precursor Sources

Are there not other chemicals besides CO and CH4 that also are contained in the precursor mix of ozone formation with its rapidly forming and degradation in the atmosphere?

Probably, but not with identified health effects at ambient levels. "The dose makes the poison."

Does the singling out of these two constituents of the ozone "cocktail" significant as push toward climate change/global warming instead of just evaluation ozone formation?

Not clear.

1.3.1.2.1 Global Methane

Again, is a teasing out/focusing on CH4 important in discussing the virtual "cocktail" of chemicals that may be associated with ozone formation/degradation?

Not relevant for human inhalation health effects.

1.3.1.2.2 International Emissions of Ozone Precursors

This section focuses on international transport of ozone precursors. What about local/state/regional transport of ozone precursors?

State/local transport is important in devising control strategies but not in evaluating health effects and setting NAAQS.

1.3.1.3.2 Biogenic Volatile Organic Compounds (VOCs)

It has been stated that biogenic VOCs and contributions are greater than anthropogenic sources (i.e., motor vehicles). Is there greater confidence in using models and remote sensing (both with relative degrees of uncertainty) to estimate biogenic ozone source contributions that vehicle emissions estimates (manufacturing vehicle emission standards and testing), in making this assessment?

I'm not sure, but these contributions probably vary with urbanicity and traffic density. We can't study a pollutant that has not been monitored and EPA holds the keys as to what gets monitored, which has been limited to what has already been declared toxic in the absence of such data. I see this as a circular situation closed to new information.

1.4 Ozone Photochemistry

With the advent of monitoring for speciated compounds including PAMS and Near-Road Monitoring (NOy), should there be further discussions about the individual chemicals gleaned from the specialized monitoring.

Not without laboratory work to determine toxicity.

1.5 Inter-Annual Variability and Longer-Term Trends in Meteorological Effects on Anthropogenic and US Background (USB) Ozone

While temperature, wind patterns, cloud cover, and precipitation are highlighted as very important variables in ozone formation, does topography play a role (such as in Birmingham where summertime pollutants are trapped in a "mountainous bowl?"

Yes, to the extent that atmospheric residence time is increased. This is the case with Los Angeles, hence their exemption from national standards. I'm not sure we know much about the "background" ozone in LA.

Are there any independent effects on formation due to relative humidity?

Water vapor plays a role in reducing photochemistry, perhaps by reducing temperature levels.

Appendix 2 Exposure to Ambient Ozone

2.3 Exposure Assessment Methods

While monitoring, including fixed, ambient monitors and personal and microenvironmental monitors are highlighted, what about remote sensing? Biological sampling in blood or tissue?

Personal monitoring is virtually nonexistent, including for ozone. Foley et al. (2003) reported that EPA considered outdoor ambient air quality from fixed ambient monitoring sites to be a "surrogate for exposure". Biological sampling is useful in clinical experiments but not in epidemiology. I'm unaware of any autopsy that identified long-term exposure to air pollution as an underlying cause of death.

2.3.2.1 Spatial Interpolation

*While attempting to quantity concentrations at locations and areas between concentration points is included under **2.3.2 Modeling**, many of these exact same methods (i.e., data averaging, IDW, and kriging) are also utilized for **Monitoring** data shortcomings.*

2.4.1 Time-Activity Data

Is it possible that ozone exposure through time-activity data may be reduced due to temperature alone, as more people tend to avoid time spent outdoors in the summers during extremely warm/hot/humid, stagnant days which are oftentimes conditions for greater ozone formation?

Not likely, because of differential use of air conditioning according to socioeconomic status, for example. Studies of the general public may be misleading, and temporal trends in the use of A/C should be considered in long-term studies, especially with regard to accountability for ozone reductions.

Miscellaneous Question(s)

Due to exposure to ozone being disproportionate for disparate groups (i.e., lower income, children), should this be an emphasis in this section, in lieu of regression analysis confounding/covariate in epidemiological studies for low(er) SES?

Yes, especially in spatial analysis.

Questions from Dr. Steven Packham

Question 1 Background Statement of Fact:

> *Evidence from controlled human exposures is sufficient to conclude with certainty that a causal relationship exists between measurable decrements in FEV1 and subjective symptoms in healthy human adults.*

Question 1: When a causal relationship is conclusive to a high degree of scientific certainty as it is in this case, should this take precedence over causal inference when drafting a NAAQS ISA?

Yes. Figure ES-3 establishes that short-term exposure to O_3 affects lung function. Figure 6 above shows a relationship between FEV_1 change and mortality, thus extending the interpretation of the FEV_1 relationship. Experimental results should take precedence over theory and statistics but care must be taken in selecting experimental subjects.

Question 2 Background Statements of Fact:

1. *The shape of the ozone induced FEV1 and subjective symptoms dose-response curve is a function of the inhaled hourly dosage rate and the cumulative dose inhaled over several hours immediately prior to the onset of the effect.*
2. *The mean cumulative dose threshold for ozone induced FEV1 and symptom effects in healthy adult humans exposed 6.6 hours to ozone concentrations from 60 to 87 ppb is estimated to be 1,362 mg. (Schelegle et al. 2009)*
3. *This is equivalent to inhaling a dose of 2,439 trillion highly reactive oxidizing molecular moieties.*
4. *Whatever the oxidative challenge of PM air pollution is to the human lung, it pales in significance to that of ozone.*
5. *The inhaled hourly dosage rate and cumulative dose thresholds appear to be lower for ozone induced FEV1 and symptom responses than those necessary for inducing clinical signs of injurious pulmonary inflammation.*
6. *Ozone induced FEV1 decrement and subjective symptoms may be species-specific protective and defensive responses and warning signs for human organisms.*
7. *Ozone exposures have been shown to stimulate peripheral mucus flow into central bronchi thereby enhancing particle transport from peripheral to central airways and mucociliary clearance of inhaled particulate matter. This beneficial dose dependent response to ozone "…is of interest since it characterizes the reaction of a primary defense mechanism essential to the protection of mucosal surfaces of the tracheobronchial tree." (Forster et al. 1987)*

Question 2: Given evidence available from controlled human exposures substantiating causal relationships with a number of physiological responses, including beneficially confounding interactions of ozone on PM clearance, should Sub-section ES4.1 Health Effects in the Draft's Executive Summary, and the entire Integrated Synthesis section of the Draft be rewritten?

Yes, for sure!

Question 3 Background Information: Figure ES-3 in the Ozone ISA External Review Draft (shown below) is adapted from the 2013 Ozone ISA which was based on eight human studies published between 1988 and 2013. The 2009 study by Schelegle et al. played a decisive role in the 2015 revision of the O3 NAAQS from 75 to 70 ppb (80 FR 65292 Oct 26, 2015).

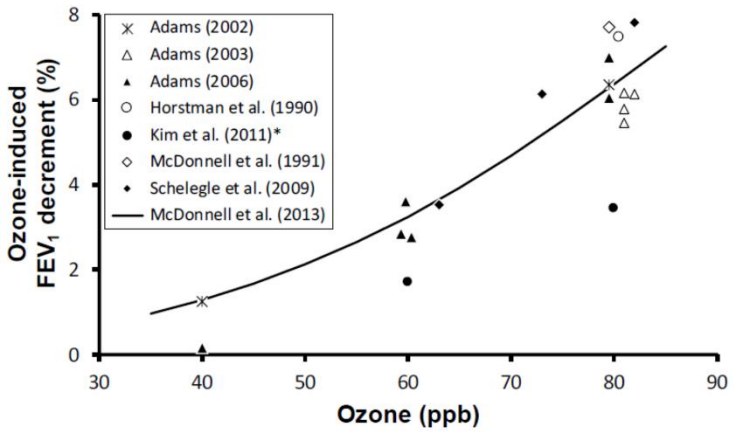


Figure ES-3 *was adapted from Figure 6-1 of 2013 Ozone ISA (U.S. EPA, 2013) which was based on studies by Adams (2006), Adams (2003), Adams (2002), Folinsbee et al. (1988), Horstman et al. (1990), Kim et al. (2011), McDonnell et al. (2013), McDonnell et al. (1991), and Schelegle et al. (2009).*

Figure 1 below (from Schelegle et al. 2009), on the other hand, depicts the actual mean accumulative doses of 31 healthy adult human subjects who completed four 6.6-hour chamber exposures to target mean O_3 concentrations of 60,70, 80, and 87 ppb. The original data presented in this way conveys critical information to toxicologists and biomedical researchers that is "lost in translation" in the concentration/risk-effect picture presented in Figure ES-3.

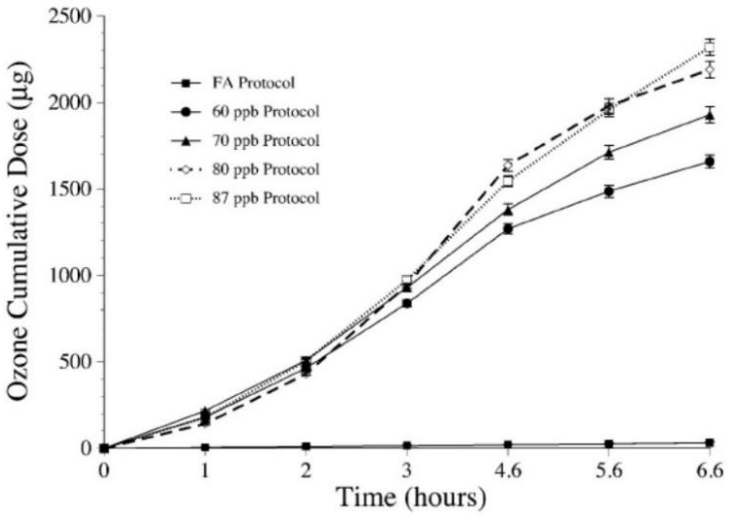
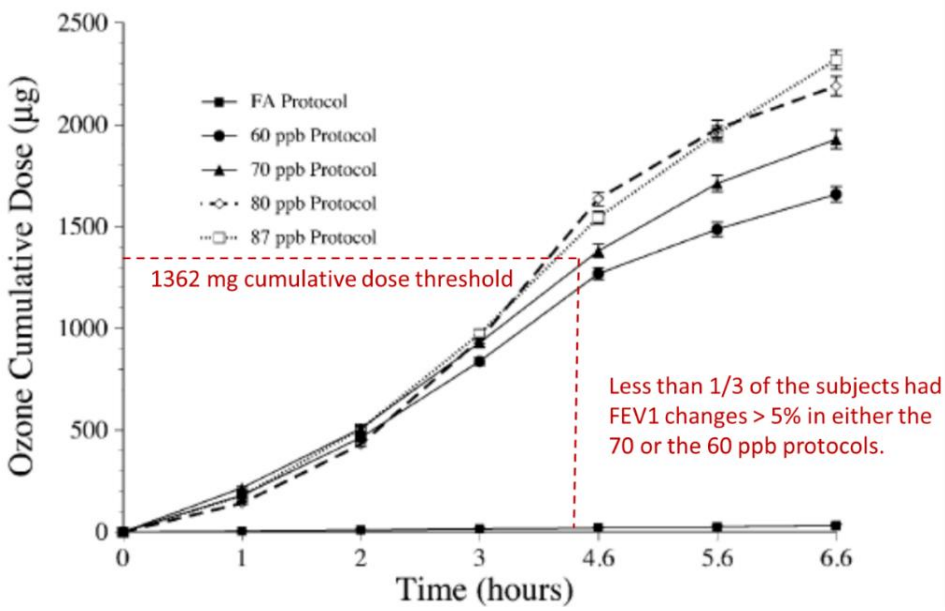


Figure 1. Diagram of mean group values for cumulative dose of ozone (micrograms) against time of exposure for each of the five protocols. Values represent means ± SEM.

To quote Schelegle et al. (2009),

> *"We were able to obtain reliable estimates of a Dose of Onset [i.e., a threshold for the FEV1 effect], using the pooled FEV1from the 80 and 87 ppb ozone exposure protocols, …but not from the pooled FEV1 data from the 60 and 70 ppb ozone exposure protocols. The inability to estimate [a threshold] using the FEV1 data from the 60 and 70 ppb ozone exposure protocols is most likely because less than one third of the subjects had changes in FEV1 greater than 5% in either of these protocols. (Emphasis added)*



Packham Figure 1. Adapted from Schelegle et al. (2009) with toxicological annotations by author, 2019.

The notable differences between Figure ES-3 compared with Packham Figure 1 are driven by how data are interpreted by different scientific disciplines. By superimposing Schelegle's descriptive conclusion-narrative onto the Sigmoid shaped dose-response curves, one sees the beginning of an increased trend of dose-response curve separation between hour 3 and hour 4: Indicative of the cumulative Dose of Onset threshold between the respective exposure protocols.

Figure ES-3 is the product of adapting (i.e., imposing) an ISA Preamble quantal risk-assessment mindset upon graded data collected from continuous response gradients characteristic of living biological organisms. The narrative associated with Figure ES-3 (found on page ES-7) is grossly misleading with respect to the epidemiologically "associated" adverse health effects and completely overlooks the confounding health benefit of enhanced PM clearance stimulated by 200 ppb ozone exposures mentioned above under Question 2 Background Statements of Fact.

The controlled human studies by Folinsbee, Adams, Horstman, Kim, McDonnell and Schelegle, and others cited below in the References and reading list, prove with absolute certainty that exposures to elevated ambient levels of O3 can cause measurable decrements in FEV1 pulmonary test results in healthy adults. These studies document that the effect of O3 on reduced FEV1 volumes is temporary, and

suggest that hourly mean ambient O3 concentrations below 70 ppb are not likely to cause FEV1 effects in most healthy adults.

Question 3 Background Statement of Facts:

Several nonmember consultants have expressed reluctance to comment on certain questions because of limited familiarity with pulmonary physiology and inhalation toxicology. Here are few facts to keep in mind.

1. *Lungs have an evolutionary history in which surfactant was key to the evolution of all air breathing species on the surface of the planet, (Daniels and Orgeig (2003.)*
2. *Antioxidant secretions from epithelial Type II cells into the liquid lining of the lungs is one of most important natural defenses the human organism has against naturally occurring ozone levels in the atmosphere near the earth's surface.*
3. *All known effects of ozone on the human respiratory system are dose dependent.*
4. *Ozone stimulation of the respiratory airways evokes a number of organism defensive and adaptive responses in humans.*
5. *Ozone alters tracheobronchial mucociliary function in humans resulting in enhanced transport and clearance of particles deposited in the peripheral air ways, (Foster, et al (1987).*
6. *Ozone is a potent oxidizing agent, (Pryor et al. (1991).*

Question 3 Overarching Conceptual Contexts: An accurate understanding of the causal dose-response relationship between ambient ozone exposure and responses elicited in the human organism opens up a number of important options that could be considered in reviewing and setting NAAQS standards and in how those standards might be used to protect, and even promote, public health. For instance, the realization that the ozone-induced FEV1 effects are temporary, reversible, and occur at a lower inhaled dose than a truly adverse health effect (such as a nonhealing, injurious inflammatory response) could be considered a tenable rationale for classifying them as natural, organism-specific margin-of-safety benchmark indicators.

Another application of hourly MSS inhalation dosage models and thresholds would be to imbed them into web and mobile platform applications for public education and development of user-friendly air quality risk management tools by the EPA. As proof of this concept's possibility, there are two air pollution exposure apps presently in the public domain: A web app http://webapp0.myairhealth.com/# and a free downloadable smartphone app https://apps.apple.com/us/app/myair-health/id790049340.

Question 3: Looking ahead, do you think toxicology, clinical human studies, and biomedical research disciplines should be given more explicit and balanced consideration in the development of the present, and future, O_3 ISAs with the objective to validate causal relationships and determine hourly inhalation dosage rates for adverse inflammatory responses in pulmonary tissues?

Yes, I whole heartedly agree. Experimental research under carefully controlled conditions (including selection of subjects can show what *can* happen under these conditions and *why*, but epidemiology is required to show actually *does* happen in the real world. Statistically significant null findings may thus be particularly important.

However, epidemiology is currently beset with several unsurmountable problems that do not extend to experimental studies:

1. Unknown personal exposures, largely because indoor air is now often dirtier than outdoor (but not for ozone).
2. Limited selection of pollutants of interest. EPA has full control over the selection of pollutants to be monitored in the ambient and makes regulatory rather scientific selections. PM constituents and ultrafines are important cases in point. Perhaps VOCs should also be considered
3. Obsolete cohorts. If EPA had commissioned a new cohort ca. 2000 instead of relying on private (ACS) data that began in 1982, the science would now be in much better shape.
4. The inability to focus on individuals with varying pre-existing conditions.

A downside here is the lack of and difficulty in conducting truly long-term or longitudinal morbidity studies, including experimental studies. I would counter this with an over-arching message: there is no convincing evidence for truly long-term effects on mortality that cannot be explained as the sum of short-term effects. Air quality and experimental research should be moved out of the regulatory arena to a purely scientific one.

Without getting into the details of Schelegle et al. (2009) I would note that variability in individual responses supports the time-series mortality model of Murray et al. based on differences in frailty. Almost all of the extant short-term epidemiology assumes uniform responses throughout the entire population, usually everyone aged 65 and over. Variability in robustness to environmental excursions is responsible for the inability to detect short-term thresholds since in a large population there will always be someone on death's doorstep (for whatever reason) who will succumb to a minor excursion. Society must thus decide on a tolerable level of risk.

Appendix A. Other relevant papers by Lipfert and colleagues not cited in the ISA

Lipfert FW, Morris SC, Wyzga RE. Daily mortality in the Philadelphia metropolitan area and size-classified particulate matter. J Air Waste Manag Assoc. 2000 Aug;50(8):1501-13.

Lipfert FW, Perry HM Jr, Miller JP, Baty JD, Wyzga RE, Carmody SE. The Washington University-EPRI Veterans' Cohort Mortality Study: preliminary results. Inhal Toxicol. 2000;12 Suppl 4:41-73.
Baxter LA, Finch SJ, Lipfert FW, Yu Q. Comparing estimates of the effects of air pollution on human mortality obtained using different regression methodologies. Risk Anal. 1997 Jun;17(3):273-8.
Lipfert FW. A critical review of studies of the association between demands for hospital services and air pollution. Environ Health Perspect. 1993 Jul;101Suppl 2:229-68. Review.
Lipfert FW, Hammerstrom T. Temporal patterns in air pollution and hospital admissions. Environ Res. 1992 Dec;59(2):374-99.

Appendix B.

B-1 Papers involving short-term associations between ozone and daily mortality in U.S. populations.

1: Murray CJ, Lipfert FW. Inferring frail life expectancies in Chicago from daily fluctuations in elderly mortality. Inhal Toxicol. 2013 Jul;25(8):461-79.
2: Ensor KB, Raun LH, Persse D. A case-crossover analysis of out-of-hospital cardiac arrest and air pollution. Circulation. 2013 Mar 19;127(11):1192-9.

3: Sacks JD, Ito K, Wilson WE, Neas LM. Impact of covariate models on the assessment of the air pollution-mortality association in a single- and multipollutant context. Am J Epidemiol. 2012 Oct 1;176(7):622-34.

4: Murray CJ, Lipfert FW. A new time-series methodology for estimatingrelationships between elderly frailty, remaining life expectancy, and ambient air quality. Inhal Toxicol. 2012;24(2):89-98.

5: Zanobetti A, Schwartz J. Ozone and survival in four cohorts with potentially predisposing diseases. Am J Respir Crit Care Med. 2011 Oct 1;184(7):836-41.

6: Peng RD, Bobb JF, Tebaldi C, McDaniel L, Bell ML, Dominici F. Toward a quantitative estimate of future heat wave mortality under global climate change. Environ Health Perspect. 2011 May;119(5):701-6.

7: Chang HH, Zhou J, Fuentes M. Impact of climate change on ambient ozone level and mortality in southeastern United States. Int J Environ Res Public Health. 2010 Jul;7(7):2866-80.

8: Murray CJ, Lipfert FW. Revisiting a population-dynamic model of air pollution and daily mortality of the elderly in Philadelphia. J Air Waste Manag Assoc. 2010 May;60(5):611-28.

9: Ren C, Williams GM, Mengersen K, Morawska L, Tong S. Temperature enhanced effects of ozone on cardiovascular mortality in 95 large US communities, 1987-2000: Assessment using the NMMAPS data. Arch Environ Occup Health. 2009 Fall;64(3):177-84.

10: Katsouyanni K, Samet JM, Anderson HR, Atkinson R, Le Tertre A, Medina S, Samoli E, Touloumi G, Burnett RT, Krewski D, Ramsay T, Dominici F, Peng RD, Schwartz J, Zanobetti A; HEI Health Review Committee. Air pollution and health: a European and North American approach (APHENA). Res Rep Health Eff Inst. 2009 Oct;(142):5-90.

11: Smith RL, Xu B, Switzer P. Reassessing the relationship between ozone and short-term mortality in U.S. urban communities. Inhal Toxicol. 2009 Sep;21 Suppl 2:37-61.

12: Ostro BD, Roth LA, Green RS, Basu R. Estimating the mortality effect of the July 2006 California heat wave. Environ Res. 2009 Jul;109(5):614-9.

14: Medina-Ramón M, Schwartz J. Who is more vulnerable to die from ozone air pollution? Epidemiology. 2008 Sep;19(5):672-9.

15: Zanobetti A, Schwartz J. Is there adaptation in the ozone mortality relationship: a multi-city case-crossover analysis. Environ Health. 2008 May 30;7:22.

16: Basu R, Feng WY, Ostro BD. Characterizing temperature and mortality in nine California counties. Epidemiology. 2008 Jan;19(1):138-45.

17: Ito K, Thurston GD, Silverman RA. Characterization of PM2.5, gaseous pollutants, and meteorological interactions in the context of time-series health effects models. J Expo Sci Environ Epidemiol. 2007 Dec;17 Suppl 2:S45-60.

18: Medina-Ramón M, Schwartz J. Temperature, temperature extremes, and mortality: a study of acclimatisation and effect modification in 50 US cities. Occup EnvironMed. 2007 Dec;64(12):827-33.

19: Bell ML, Kim JY, Dominici F. Potential confounding of particulate matter on the short-term association between ozone and mortality in multisite time-series studies. Environ Health Perspect. 2007 Nov;115(11):1591-5.

20: Sarnat SE, Suh HH, Coull BA, Schwartz J, Stone PH, Gold DR. Ambient particulate air pollution and cardiac arrhythmia in a panel of older adults in Steubenville, Ohio. Occup Environ Med. 2006 Oct;63(10):700-6. 4.

21: Ostro BD, Tran H, Levy JI. The health benefits of reduced tropospheric ozone in California. J Air Waste Manag Assoc. 2006 Jul;56(7):1007-21.

22: Zeger SL, Irizarry R, Peng RD. On time series analysis of public health and biomedical data. Annu Rev Public Health. 2006;27:57-79.
23: Ito K, De Leon SF, Lippmann M. Associations between ozone and daily mortality: analysis and meta-analysis. Epidemiology. 2005 Jul;16(4):446-57.

24: Bell ML, Dominici F, Samet JM. A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. Epidemiology. 2005 Jul;16(4):436-45.

25: Schwartz J. How sensitive is the association between ozone and daily deaths to control for temperature? Am J Respir Crit Care Med. 2005 Mar 15;171(6):627-31.

26: Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F. Ozone and short-term mortality in 95 US urban communities, 1987-2000. JAMA. 2004 Nov 17;292(19):2372-8.

27: Schwartz J. Is the association of airborne particles with daily deaths confounded by gaseous air pollutants? An approach to control by matching. Environ Health Perspect. 2004 Apr;112(5):557-61.

28: Moolgavkar SH. Air pollution and daily mortality in two U.S. counties: season-specific analyses and exposure-response relationships. Inhal Toxicol. 2003 Aug;15(9):877-907.

29: Thurston GD, Ito K. Epidemiological studies of acute ozone exposures and mortality. J Expo Anal Environ Epidemiol. 2001 Jul-Aug;11(4):286-94. Review.
30: Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994. N Engl J Med. 2000 Dec 14;343(24):1742-9.

31: Ostro BD, Broadwin R, Lipsett MJ. Coarse and fine particles and daily mortality in the Coachella Valley, California: a follow-up study. J Expo Anal Environ Epidem. 2000;10(5):412-9.
32: Moolgavkar SH. Air pollution and daily mortality in three U.S. counties. Environ Health Perspect. 2000 Aug;108(8):777-84.

33: Lippmann M, Ito K, Nádas A, Burnett RT. Association of particulate matter components with daily mortality and morbidity in urban populations. Res Rep Health Eff Inst. 2000 Aug;(95):5-72,

34: Mar TF, Norris GA, Koenig JQ, Larson TV. Associations between air pollution and mortality in Phoenix, 1995-1997. Environ Health Perspect. 2000 Apr;108(4):347-53.

35: Ito K, Thurston GD. Daily PM10/mortality associations: an investigation of at-risk subpopulations. J Expo Anal Environ Epidemiol. 1996 Jan-Mar;6(1):79-95.

36: Moolgavkar SH, Luebeck EG, Hall TA, Anderson EL. Air pollution and daily mortality in Philadelphia. Epidemiology. 1995 Sep;6(5):476-84.

37: Ostro B. Fine particulate air pollution and mortality in two Southern California counties. Environ Res. 1995 Aug;70(2):98-104.

38: Dockery DW, Schwartz J, Spengler JD. Air pollution and daily mortality: associations with particulates and acid aerosols. Environ Res. 1992 Dec;59(2):362-73.

39: Pope CA 3rd, Schwartz J, Ransom MR. Daily mortality and PM10 pollution in Utah Valley. Arch Environ Health. 1992 May-Jun;47(3):211-7.

40: Ramlow JM, Kuller LH. Effects of the summer heat wave of 1988 on daily mortality in Allegheny County, PA. Public Health Rep. 1990 May-Jun;105(3):283-9.

41: Liu T, Zeng W, Lin H, Rutherford S, Xiao J, Li X, Li Z, Qian Z, Feng B, Ma W. Tempo-Spatial Variations of Ambient Ozone-Mortality Associations in the USA: Results from the NMMAPS Data. Int J Environ Res Public Health. 2016 Aug 26;13(9).pii: E851.

42: Hao Y, Balluz L, Strosnider H, Wen XJ, Li C, Qualters JR. Ozone, Fine Particulate Matter, and Chronic Lower Respiratory Disease Mortality in the UnitedStates. Am J Respir Crit Care Med. 2015 Aug 1;192(3):337-41.

43: Madrigano J, Jack D, Anderson GB, Bell ML, Kinney PL. Temperature, ozone, and mortality in urban and non-urban counties in the northeastern United States. Envir Health. 2015 7;14:3.

44: Jhun I, Fann N, Zanobetti A, Hubbell B. Effect modification of ozone-related mortality risks by temperature in 97 US cities. Environ Int. 2014 Dec;73:128-34.

45 Bell ML, Zanobetti A, Dominici F. Who is more affected by ozone pollution? A systematic review and meta-analysis. Am J Epidemiol. 2014 Jul 1;180(1):15-28.

46: Di Q, Dai L, Wang Y, Zanobetti A, Choirat C, Schwartz JD, Dominici F. Association of Short-term Exposure to Air Pollution With Mortality in Older Adults. JAMA. 2017 318(24):2446-2456.
47: Young SS, Smith RL, Lopiano KK. Air quality and acute deaths in California, 2000-2012. Regul Toxicol Pharmacol. 2017 Aug;88:173-184.

B-2 Papers involving long-term associations between ozone and daily mortality

1. Hvidtfeldt UA, Sørensen M, Geels C et al. Long-term residential exposure to PM(2.5), PM(10), black carbon, NO(2), and ozone and mortality in a Danish cohort. Environ Int. 2019 Feb;123:265-272.

2. Sanyal S, Rochereau T, Maesano CN et al.. Long-Term Effect of Outdoor Air Pollution on Mortality and Morbidity: A 12-Year Follow-Up Study for Metropolitan France. Int J Environ Res Public Health. 2018 Nov 8;15(11). pii: E2487.

3. Lipfert FW, Wyzga RE. Revisiting the Veterans Cohort Mortality Study: New results and synthesis. J Air Waste Manag Assoc. 2018 Nov;68(11):1248-1268.

4. Chen K, Wolf K, Breitner S, Gasparrini A et al. Two-way effect modifications of air pollution and air temperature on total natural and cardiovascular mortality in eight European urban areas. Environ Int. 2018 Jul;116:186-196.

5. Nuvolone D, Petri D, Voller F. The effects of ozone on human health. Environ Sci Pollut Res Int. 2018 Mar;25(9):8074-8088.

6. Cakmak S, Hebbern C, Pinault L et al. Associations between long-term PM(2.5) and ozone exposure and mortality in the Canadian Census Health and Environment Cohort (CANCHEC), by spatial synoptic classification zone. Environ Int. 2018 Feb;111:200-211.

7. Robertson S, Miller MR. Ambient air pollution and thrombosis. Part Fibre Toxicol. 2018 Jan 3;15(1):1.

8. Bourdrel T, Bind MA, Béjot Y, Morel O, Argacha JF. Cardiovascular effects of air pollution. Arch Cardiovasc Dis. 2017 Nov;110(11):634-642.

9. Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Choirat C, Dominici F, Schwartz JD. Air Pollution and Mortality in the Medicare Population. N Engl J Med. 2017 Jun 29;376(26):2513-2522.

10. Cakmak S, Hebbern C, Vanos J, Crouse DL, Burnett R. Ozone exposure and cardiovascular-related mortality in the Canadian Census Health and Environment Cohort (CANCHEC) by spatial synoptic classification zone. Environ Pollut. 2016

11. Turner MC, Jerrett M, Pope CA 3rd et al. Long-Term Ozone Exposure and Mortality in a Large Prospective Study. Am J Respir Crit Care Med. 2016 May 15;193(10):1134-42.

12. Yang WS, Zhao H, Wang X, Deng Q, Fan WY, Wang L. An evidence-based assessment for the association between long-term exposure to outdoor air pollution and the risk of lung cancer. Eur J Cancer Prev. 2016 May;25(3):163-72.

13. Thurston GD, Ahn J, Cromar KR et al. Ambient Particulate Matter Air Pollution Exposure and Mortality in the NIH-AARP Diet and Health Cohort. Environ Health Perspect. 2016 Apr;124(4):484-90.

14. Green R, Broadwin R, Malig B et al. Long- and Short-term Exposure to Air Pollution and Inflammatory/Hemostatic Markers in Midlife Women. Epidemiology. 2016 Mar;27(2):211-20.

15. Atkinson RW, Butland BK, Dimitroulopoulou C et al. Long-term exposure to ambient ozone and mortality: a quantitative systematic review and meta-analysis of evidence from cohort studies. BMJ Open. 2016 Feb

16. Crouse DL, Peters PA, Hystad P et al. Ambient PM2.5, O_3, and NO_2 Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ Health Perspect. 2015 Nov;123(11):1180-6.

17. Hao Y, Balluz L, Strosnider H et al. Ozone, Fine Particulate Matter, and Chronic Lower Respiratory Disease Mortality in the United States. Am J Respir Crit Care Med. 2015 Aug 1;192(3):337-41.

18. Prueitt RL, Lynch HN, Zu K, Sax SN, Venditti FJ, Goodman JE. Weight-of-evidence evaluation of long-term ozone exposure and cardiovascular effects. Crit Rev Toxicol. 2014 Oct;44(9):791-822.

19. Carey IM, Atkinson RW, Kent AJ et al. Mortality associations with long-term exposure to outdoor air pollution in a national English cohort. Am J Respir Crit Care Med. 2013 Jun 1;187(11):1226-33.

20. Atkinson RW, Carey IM, Kent AJ et al. Long-term exposure to outdoor air pollution and incidence of cardiovascular diseases. Epidemiology. 2013 Jan;24(1):44-53.

21. Wang XY, Hu W, Tong S. Long-term exposure to gaseous air pollutants and cardio-respiratory mortality in Brisbane, Australia. Geospat Health. 2009 May;3(2):257-63.

22. Jerrett M, Burnett RT, Pope CA 3rd et al. Long-term ozone exposure and mortality. N Engl J Med. 2009 Mar 12;360(11):1085-95.

23. Fischer P, Hoek G, Brunekreef B, Verhoeff A, van Wijnen J. Air pollution and mortality in The Netherlands: are the elderly more at risk? Eur Respir J Suppl. 2003 May;40:34s-38s.

24. Mills PK, Abbey D, Beeson WL, Petersen F. Ambient air pollution and cancer in California Seventh-day Adventists. Arch Environ Health. 1991 Sep-Oct;46(5):271-80.

25. Abbey DE, Mills PK, Petersen FF, Beeson WL. Long-term ambient concentrations of total suspended particulates and oxidants as related to incidence of chronic disease in California Seventh-Day Adventists. Environ Health Perspect. 1991 Aug;94:43-50.

26. Jerrett M, Burnett RT, Beckerman BS et al. Spatial analysis of air pollution and mortality in California. Am J Respir Crit Care Med. 2013 Sep 1;188(5):593-9.

27. Jerrett M, Burnett RT, Ma R, Pope CA 3rd et al. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology. 2005 Nov;16(6):727-36.

Abstracts of key papers

Frank R, Tankersley C. Air pollution and daily mortality: a hypothesis concerning the role of impaired homeostasis. Environ Health Perspect. 2002 Jan;110(1):61-5.
We propose a hypothesis to explain the association between daily fluctuations in ambient air pollution, especially airborne particles, and death rates that can be tested in an experimental model. The association between airborne particulates and mortality has been observed internationally across cities with differing sources of pollution, climates, and demographies and has involved chiefly individuals with advanced chronic illnesses and the elderly. As these individuals lose the capacity to maintain stable, optimal internal environments (i.e., as their homeostatic capacity declines), they become increasingly vulnerable to external stress. To model homeostatic capacity for predicting this vulnerability, a variety of regulated physiologic variables may be monitored prospectively. They include the maintenance of deep body temperature and heart rate, as well as the circadian oscillations around these set-points. Examples are provided of the disruptive changes shown by these variables in inbred mice as the animals approach death. We consider briefly the implications that the hypothesis may hold for several epidemiologic issues, including the degree of prematurity of the deaths, the unlikelihood of a threshold effect, and the role that coarse, non-combustive particles may play in the association.

Lipfert FW. Long-term associations of morbidity with air pollution: A catalog and synthesis. J Air Waste Manag Assoc. 2018 Jan;68(1):12-28.

ABSTRACT: I searched the National Institutes of Health MEDLINE database through January 2017 for long-term studies of morbidity and air pollution and cataloged them with respect to cardiovascular, respiratory, cancer, diabetes, hospitalization, neurological, and pregnancy-birth endpoints. The catalog is presented as an online appendix. associations with PM2.5 (particulate matter with an aerodynamic diameter <2.5 μm), PM10 (PM with an aerodynamic diameter <10 μm), and nitrogen dioxide (NO2) were evaluated most frequently among the 417 ambient air quality studies identified. Associations with total suspended particles (TSP), carbon, ozone, sulfur, vehicular traffic, radon, and indoor air quality were also reported. I evaluated each study in terms of pollutant significance (yes, no), duration of exposure, and publication date. I found statistically significant pollutant relationships ($P < 0.05$) in 224 studies; 220 studies indicated adverse effects. Among 795 individual pollutant effect estimates, 396 are statistically significant. Pollutant associations with cardiovascular indicators, lung function, respiratory

symptoms, and low birth weight are more likely to be significant than with disease incidence, heart attacks, diabetes, or neurological endpoints. Elemental carbon (EC), traffic, and PM2.5 are most likely to be significant for cardiovascular outcomes; TSP, EC, and ozone (O3) for respiratory outcomes; NO2 for neurological outcomes; and PM10 for birth/pregnancy outcomes. Durations of exposure range from 60 days to 35 yr, but I found no consistent relationships with the likelihood of statistical significance. Respiratory studies began ca. 1975; studies of diabetes, cardiovascular, and neurological effects increased after about 2005. I found 72 studies of occupational air pollution exposures; 40 reported statistically significant adverse health effects, especially for respiratory conditions. I conclude that the aggregate of these studies supports the existence of nonlethal physiological effects of various pollutants, more so for non-life-threatening endpoints and for noncriteria pollutants (TSP, EC, PM2.5 metals). However, most studies were cross-sectional analyses over limited time spans with no consideration of lag or disease latency. Further longitudinal studies are thus needed to investigate the progress of disease incidence in association with air pollution exposure.

IMPLICATIONS: Relationships of air pollution with excess mortality are better known than with long-term antecedent morbidity. I cataloged 489 studies of cardiovascular, respiratory, cancer, and neurological effects, diabetes, and birth outcomes with respect to 12 air pollutants. About half of the studies reported statistically significant relationships, more frequently with noncriteria than with criteria pollutants. Indoor and cumulative exposures, coarse or ultrafine particles, and organic carbon were seldom considered. Significant relationships were more likely with less-severe endpoints such as blood pressure, lung function, or respiratory symptoms than with incidence of cancer, chronic obstructive pulmonary disease (COPD), heart failure, or diabetes. Most long-term studies are based on spatial relationships; longitudinal studies are needed to link the progression of pollution-related morbidity to mortality, especially for the cardiovascular system.

Dr. D. Warner North, NorthWorks

I begin by reiterating three themes from my earlier response on the Policy Assessment for Particulate Matter.

I. **Causality.** I do not endorse, and am troubled by, EPA's existing Causality framework, Table II, "Weight of evidence for causality determinations," introduced in the Preface at pages lxxix-lxxxx, then again at ES.4 and IS.1.2.4, page IS-6, with determinations set forth in Figure ES-2 and Table IS-1, page IS-7. I support the criticisms of this framework in the two peer-reviewed publications of Dr. Tony Cox referenced in his Questions, ("Modernizing the Bradford Hill criteria," in *Crit. Rev. Tox.* (2018) and "Improving causal determination," in *Global Epidemiology*, 2019). I judge that to deal with these criticisms, much of the ISA would need to be greatly revised – essentially, rewritten. I question whether it is possible to do that on the existing schedule. Ozone causes adverse health impacts on human health at levels at or above 5 parts per million, which is the level listed as "immediately dangerous to life and health" (IDLH) by NIOSH: https://www.cdc.gov/niosh/idlh/10028156.html. The focus of CASAC's review should be on whether there is clear evidence that ozone is a significant partial causal contributor to adverse human health impacts at levels nearly two orders of magnitude lower, that is, at or below 120 parts per billion (the 1979-1997 one hour standard), and especially, in the range between 70 ppb, the existing 8-hour standard, and 60 ppb, a lower standard that might be considered as providing increased protection of public health including susceptible groups, with "an adequate margin of safety."

II. **Concentration- Response relationship.** For both epidemiological studies and human clinical studies, much of the data showing responses above background occurs at levels above this range of 60 to 70 ppb. How should the extrapolation be made to predict changes in mortality, morbidity such as respiratory and cardiovascular disease, or clinical changes such as FEV1 as a measure of lung function? For measures such as lung function, how should responses in healthy volunteers be used to predict adverse health impacts in children, elderly, and members of sensitive groups? Use of linear-to-zero extrapolation should be compared to alternatives, with attention to biological plausibility, variability, and uncertainty.

III. **Regional issues resulting in extreme exposures, such as wildfires.** As a resident of northern California, I am concerned that a large number of the "red circles" in Figure ES-1, page ES-2, (also, Figure 1-8, Appendix 1, p.1-44) of violations of the standard (since it was lowered from 120 ppb in 1997) of 76 to 112 ppb occur in central California. (There are many "orange circles" in California as well, but the majority of orange circles corresponding to small exceedances of the 70 ppb current primary standard are in urban centers in states east of California.) It is not clear in the text whether these "design values" include wildfire episodes, or whether wildfire episodes with even higher ozone observations were removed as "natural disaster" exceptions. Figure 1-12 page 1-48 shows the large number of design values in California, plus a few in other western states, with exceedances of the 70 ppb MDA8 standard and high W126 sigmoidally weighted sums of hourly O3 concentrations.

> I urge that CASAC concern itself with wildfire-related ozone exposures and how these high exposures may be reduced by actions that EPA might take, or might advocate to other federal agencies and state agencies that have responsibilities affecting the occurrence and extent of wildfires. California is experiencing severe wildfires in 2019 as it has several previous years. These wildfires contribute substantially to violations of the existing 8-hour ozone standard of 70 ppb, but under EPA's rules, wildfire exposures may be excluded both from regulation and from including these episodes as data because wildfires are claimed to be "natural disasters." I urge CASAC to tell EPA that the threat of ozone exposure from wildfires should be described in detail in the ISA, and that strategies to reduce wildfires and consequent high PM and ozone levels should be discussed so that the public may be informed on this important cause of NAAQS violations.

More on Theme III, Wildfires and Ozone: Why is there not more discussion of this important source of ozone exposure in the ISA?

Wildfires are mentioned on line 5 of page ES-3 as contributing to "ground-level US background (USB) ozone." But the paragraph, lines 1-17, is quite inadequate in portraying how ozone from a large wildfire can contribute to ozone levels of 70 ppb and above, lasting for multiple days and affecting large areas such as those portions of central California with the red circles in Figure ES-1. And considering ozone from wildfires as "background" seems inappropriate, since the occurrence of wildfires is strongly influenced by human activities, especially land use and vegetation management.

Wildfires are mentioned on line 26 of page IS-13, and then again at IS-14 line 23-24:

> Wildfires have been estimated to contribute a few ppb to seasonal mean ozone concentrations in the U.S., but episodic contributions may be as high as 30 ppb (Section 1.3.1.2). However, estimates of the magnitude of ozone formation from wildfires is highly uncertain with some work showing large overpredictions of modeled wildfire contributions (Section 1.3.1.3). (Note: here is a rare example of a grammatical error in the ISA: "estimates … is")

The statement that "episodic contributions" can be as high as 30 ppb" and that ozone formation "can continue for up to five days following a wildfire event" (page 1-21, lines 38-41) should motivate attention. 30 ppb is a large fraction of a standard now at 70 ppb for a 4th highest maximum 8 hour average in a three year period. What about the magnitude of the uncertainties and the overprediction of modeled wildfire contributions? The reader might conclude that wildfires are not so important as a source of ozone. I believe more disclosure in the ISA would be useful for clarification.

IS Section 1.8, "U.S Background Ozone Concentrations" mentions wildfires as contributing to background, "USB":

> Quantification of USB ozone on days when MDA8 ozone concentrations exceed 70 ppb is more relevant to understanding USB ozone contributions at the upper end of the distribution than are seasonal mean USB ozone estimates because USB varies daily and is a function of season, meteorology, and elevation (Jaffe et al., 2018). (1-49, lines 5-9.)

But the reader is not told that large wildfires can last for weeks, with smoke plumes filled with NOx and VOC transforming into ozone at the "red circle " level in Figure ES-1, and these wildfires happen frequently in California with the plume blowing into the Central Valley.

I now look for sections 1.3.1.2 and 1.3.1.3 in Appendix 1. The section that **should** have been referenced in the Appendix 1 text is 1.3.1.3.3," Landscape Fires," p. 1-21 and 22. I read Jaffe and Wigder (2012), "Ozone production from wildfires: A critical review," and Jaffe et al. (2018), "Review, scientific assessment of background ozone over the U.S.: Implications for air quality management." Clearly, wildfires are a significant contributor, especially when meteorological conditions such as Santa Ana winds lead to large fire episodes. This pattern has happened for many decades in California, and the past three years – 2017 to 2019 - have been particularly severe.

I also found the Exceptional Events Rule, signed by EPA September 16, 2016: https://www.epa.gov/air-quality-analysis/final-guidance-preparation-exceptional-events-demonstrations-wildfire-events. Under this rule, a state can ask that ozone levels during wildfire episodes not be used for regulatory purposes. Violations of the NAAQS do not count, and these exceedances may be deleted from the data bases. Especially the latter seems to me highly inappropriate. I believe EPA should be calling attention to these exceedances of NAAQS as threats to public health, and NOT excluding them as a part of "natural background!"[1]

Figure A2-1 of this EPA 2016 document shows the increase in ozone over time versus sources of VOC and NOx for several wildfires. In one case the range of projected ozone increase was 60-85 ppb. The two others were in the 8-22 and 24-60 ppb ranges. Figure A2-2 shows the plumes, with ozone increases of 5 to 30 ppb over large area. A good discussion of the complex calculation of how O_3 is formed from VOC and NOX in wildfire plumes is in Jaffe et al. (2018), page 16 of 30. Observations and modeling for 7 urban sites in the western U.S. gave results from negative to 33 ppb, including days with MDA8 values over 70 ppb.

Figure 3 from Jaffe et al. (2018), page 8 of 30, shows observed design values (4^{th} highest MDA8 O3 levels) averaged over 2010-2014, and modeled North American background with North American anthropogenic emissions zeroed out. For most of the American West, the resulting calculation of background levels is in the 61 to 70 ppb range. Natural background - wildfires plus lightning plus pollution from Asia can take the ozone exposure level very close to the current MDA8 standard of 70 ppb! Not only wildfires, but also other aspects of so-called natural background ("USB") are important! The details in Jaffee et al. (2018) are persuasive that so-called natural background varies considerably, and these variations and the uncertainties (of the order of 10 ppb for seasonal average and higher for MDA8 average concentrations; p. 1-53, lines 27-29) ought to be considered carefully, especially in considering how much actions to reduce ozone exposure might actually accomplish in helping areas not now meeting the MDA8 standard to achieve compliance. Yes, this 2016 EPA Rule can help states (e.g., California) achieve compliance. But the exposure above the standard is still happening, and any

[1] "EE[exceptional event] influenced data can be excluded from the design value calculation if they are identified by the state agency and supported by evidence, which is then evaluated and approved by the EPA. Thus, excluding high O_3 caused by exceptional events may allow an area to be designated in attainment of the NAAQS. … the impact appears to be greatest in the western states where wildfires tend to be greater…" Quote from Jaffe et al, 2018, page 3 of 30.

consequent public health impacts from exposure above the standard are not avoided by EPA's Exceptional Event Exclusion rule.

The Kincaid Fire has burned over 77,700 acres, mostly in Sonoma County in northern California, with full containment achieved November 6, two weeks after the fire started October 23. Fires in southern California have burned many thousands of acres more in the same time period. Many red circles/triangles in the Central Valley might result from these fires. (See Figure ES-1, Figure 1-8, Figure 1-12, and text page 1-56 line 38 to I-59 line 2 plus Figures 1-13a and 14a.) These relationships should be explained to the public in the ISA, and not left to the 2016 EPA Rule document, which I could not find referenced in the ISA. (The Jaffe et al. 2018 paper is referenced in the ISA, in both IS and extensively in Appendix 1.) Daniel Jaffe was listed among the authors, contributors, and reviewers in the ISA Preface. Jaffee et al. (2018) describes the EPA guidance on event exclusion in considerable detail. I am not aware how many California fire events have been excluded, but I expect that wildfires such as Tubbs, 2017, Camp, 2018, and Kincaid, 2019, in northern California, and large fires in these same years in southern California, would qualify. Even more of the circles and triangles in the California Central Valley in the equivalents of Figures 1-8 and 1-12 with 2017-2019 data might be red as the result of these large fires.

Other Overall Comments on the ISA. It is a huge document, over 1400 pages. It is generally well written, and the organization is such that a reader can navigate readily from the summary statements to the supporting detail and references. There are remarkably few typos and similar errors. I commend Appendix 1 as excellent support for the ES and IS material. I thought Appendix 2 was overly detailed, with insufficient emphasis on the differences between human exposure indoors, where most people spend most of their time, and ambient levels outdoors. I found weakness with the discussion of inflammation in clinical studies at near-ambient exposure levels in Appendix 3, and I am dissatisfied generally with the discussions of causality and confounding. As far as I can judge, most of the relevant clinical and epidemiological studies are included – but I and others have found other relevant studies not included. There is a deficit of published papers on interpretation of the available studies, especially on uncertainty, variability and severity of health effects. Discussion of inflammation is an example. Non-experts have difficulty understanding the importance of the range of biomarkers for assessing the degree to which these biomarkers indicate a public health impact deserving of protection under the language of the Clean Air Act. I did not find any references to the journal *Risk Analysis*, for which I am an area editor, and which publishes many papers on air pollution health risks. I expect other journals may be neglected as well.

Because there were no questions on this material and my time was limited, I did not read Appendix 8 and the ES and IS sections on welfare relevant to the secondary standard for ozone.

Response to Questions from Dr. Tony Cox

This set of questions, ten pages in length, deals mostly with the causality determination framework developed previously by EPA and used as the organizing framework for this ISA. I am in support of Dr. Cox's criticisms as described in his two referenced papers in *Critical Reviews in Toxicology* and *Global Epidemiology*. I believe the Bradford Hill criteria are out of date, and in particular, strong association is not evidence for causation. But I think emphasizing so heavily the weight of evidence using a five-part hierarchy for causal determination may result in neglect of other areas important for CASAC's review. I

urge attention to my Theme II, the shape of the concentration response (C-R) relationship. It seems to me an important issue whether observed mild, apparently reversible effects such as changes in FEV1 (forced expiratory volume in one second) seen in healthy young exercising subjects imply a potential for adverse health effects in the general population. What are the adverse health effects, and how well do FEV1 changes predict them? What is the C-R relationship, not just for FEV1 changes, but for adverse health impacts that are persistent and perhaps cumulative over time, such as scarring of lung tissue so that lung function is permanently lost?

Overarching Questions 1-4: Is the ISA *clear*? I do not find it so to me. Perhaps others would find it so. Is the ISA *sound*? Not in my judgment, because there is little exploration of confounding and too much reliance on strength of association in the Bradford Hill criteria. Is the ISA *scientific*? The classification in the causal categories are judgment calls. Different people might make them differently based on the same supporting evidence. Is the ISA *policy-relevant*? I believe CASAC is asked to make judgment calls, but this causality framework is not the proper framework to do so. The judgment should be on the change in number people whose health is protected with an adequate margin of safety, and not on the choice among the five causality categories.

Specific Questions:

1. I think this is a clear **NO**. CASAC should be seeking to evaluate manipulative or interventional causation, that is, determining how many people might be added or subtracted from having their health protected with an adequate margin of safety by a change in the primary NAAQS standard. Who will be protected at a 60 ppb MDA8 O_3 standard that was not protected at the current 70 ppb MDA8 standard? How is this number to be estimated? What is the supporting evidence and chain of logic used as the basis for the estimate? What is the uncertainty in the estimate, and what are the sources of this uncertainty? Dr. Cox's *Crit. Rev. Tox.* paper has on the top of p.19, top of col. 2, some excellent insight on the difference between Bradford Hill and manipulative causality. The col. 2 paragraphs following motivate replacing EPA's five part causal categories with a revised framework. But I doubt that a replacement framework can be developed and implemented by EPA in the current cycle for ISA revision without extending the completion data for the NAAS review process.

   O_3 is clearly dangerous to health at high levels. It is much less clear how much risk to public health O_3 poses at current ambient levels. It is CASAC's job to advise the EPA Administrator on whether the existing NAAQS should be revised.

2. (a). No. I agree that the terms are not clearly and unambiguously defined. (b). No. Other factors should be considered, and positive association is not the same as necessary or sufficient. One needs to think in terms of partial causation through examination of multiple factors. Dr. Cox's example in *Crit. Rev. Tox.*, page 3-4 shows how this may be done. (c). No. I do not find that "causal" as used in the ISA implies "a manipulative causal relationship." (d,e,f). yes, no, no. But it seems inappropriate to dispute causality at high O_3 exposure levels such as 5 ppm. CASAC needs to address the extent to which ozone exposure causes health and welfare impairment at ambient levels occurring now and in the future in the United States. There are not obvious, clear answers. Judgment is needed about the uncertainties: how many people might be impaired, and to what extent? (g). No. I do not believe the determination represented

by the entries in Table ES-1 can be defended as correct. I think the Administrator, CASAC, and we who are advising CASAC could argue ad nausium about causality, and we would be better off trying to address how to estimate the extent of health response and the severity of the health response for exposures in the 60-70 ppb range. (h). no. See Dr. Cox's example, text in (b). above. (i). no. Let's use probabilities instead of seeking yes-no answers. (j). No. Sensitive subpopulations must be considered. (k). I must respond that the lines between the categories are unclear to me. I cannot attest to mutually exclusivity. Again, I think the level of ozone exposure is critical. CASAC should evaluate the significance to health (and welfare) of ozone exposures in the 60-100 ppb range, in the context of other factors. Causality for adverse health effects seems well established at higher levels in the occupational health literature, for example, the Ozone Material Safety Data Sheets. See https://www.cdc.gov/niosh/idlh/10028156html; www.amsbio.com/images/featureareas/ozilla/Ozilla-MSDS.pdf. Other sites include: https://www.ozonesolutions.com/knowledge-center/ozone-safety.html.[2] (l). I am not sure I care about "collectively exhaustive." I do not think the framework is useful. (m). I prefer probability statements and partial causation with multiple factors as in Dr. Cox's example. I do not like implied "bright lines," such as greater than 50% means likely. Such conventions need to be agreed to among the users. I do not believe that is the case here.

3. (a).(i) I am concerned that EPA's selection process is leaving out studies with negative findings for ozone. This is evidence against the ISA study selection process being comprehensive, trustworthy, and unbiased. (ii). Including Moore (2008) but not Moore (2012) is further evidence of weakness in study. (iii and iv). I will leave to others specifics on what should be included or excluded. It is my impression that there is a lot of relevant literature on interpretation/evaluation of studies that ought to be included and is not included. (See my earlier general comments on the ISA.) (v). Not clear to me. (vi). I am disappointed that the Executive Summary focuses on the causality determination, rather than a common sense summary of the relevant science including discussion of modeling assumptions, uncertainties and variability in ozone exposure. I did not find the Annex to Appendix 6 useful as a guide to "what should be done." It seemed like a defense of EPA practice as best practice, and I disagree, especially on the use of the Bradford Hill criteria. (b). As I have stated above, EPA's framework does not discuss interventions. How exposures will change with new standards is buried in the details on the draft Policy Assessment for PM, and I expect similar problems with the upcoming draft PA for ozone and related photochemical oxidants. I would like to see possible interventions made explicit at the regional level and evaluated in terms of how much these interventions would reduce exposure, and then, what impacts these would have on health and welfare measures, with results in probabilistic form. Anne Smith's 2019 paper in *Risk Analysis* demonstrates how such an evaluation might be done. On the last sentence of part (b), I

[2] When I was in my teens I used to experiment with a chemistry set, which was dangerous, and also with a Tesla spark coil. Here is what the website http://donklipstein.com/tcsafe.html says about ozone from a Tesla coil: "The sparks and intense corona easily produced by Tesla coils can produce ozone. Ozone is bad to breathe since it can corrode lung tissue. And if you are going to breathe it a few hours a day, it can be unhealthful to breathe even if the concentration is too weak to smell. Ozone can also oxidize some rubber objects. If you are having corona, you should operate the Tesla coil only in a very well ventilated area, or expose yourself to the ozone for no more than a few minutes a day."

judge that Dr. Cox's concern is well justified. I share it, and have serious reservations about basing a risk assessment for health effects on the causal determinations from the framework EPA has used in this draft ISA. I have difficulty interpreting what "likely to be causal" means in connection with possible confounding. What I find informative in Table 3-3 (and other Tables like it) is the summary of ozone levels used in the studies. In my judgment, results obtained at levels of 500 ppb (0.5 ppm) and above have little relevance unless the biological mechanism(s) involved apply at much lower levels, such as 100 ppb. (c). I read the Kai et al. (2018) paper. It demonstrates that confounding by temperature can lead to modification of mortality estimates, and that sophisticated non-linear methods may be needed for both extremes of cold and heat. For the latter, extent of air conditioning can be important, and I did not find that included in comparing US and European cities. I did not find Kai et al. (2018) in the references in ISA Appendix 3, and that concerns me a great deal. Issues of how to deal with confounding by temperature, extent of air conditioning in homes and workplaces, and socioeconomic status – interrelated factors that will differ by location – need to be carefully evaluated in order to get good estimates of mortality and morbidity responses. EPA in this ozone ISA seems far behind evolving "best practices" in how to do such analysis. I did not read through the Tétreault et al. study, but I perceive that the kind of discussion needed on confounding was not present in the ISA, just a judgment of "likely to be causal." And I do not find such judgments useful in the absence of detailed discussion of possible confounding (d). Estimating exposure "according to the centroid of the postal code" is quite crude, and misses all the subtleties whether children are outdoors, indoors in air conditioned space, or indoors in non-air conditioned space, and the extent to which they are exercising in these environments. It also may miss exposure to materials the aggravate asthma, such as pet and cockroach dander. Dr. Cox's quotes from the study persuade me that this study has serious problems of potential confounding. (e). I did not read the sensitivity analysis associated with the Tétrealt et al study. With such crude and aggregated data, it is not clear to me what sensitivity analysis would be useful. Continuing from (d), from the description Dr. Cox has given, I would be inclined to discount the Tétreault et al. study as providing "key evidence." (f). No, I find the discussions of study quality rather superficial, It is clear that some of the recent studies have initiatives for investigating confounding factors. Much more is needed, in the context of learning how to do better estimation of the C-R relationship in the low exposure range. (g). I did not find much discussion on confounding, measurement error, and unverified modeling assumptions. The discussions I did find were often superficial. In a few cases I went to the cited paper and found interesting discussion, similar to that in Kai et al.

4. I read the full text of the Michaudel et al. 2016 paper. Activation of the NLRP3 inflammazome seems important. So I went to the text to find the exposure level. I think there was a misprint: the text reads,

> Chronic ozone exposure such as twice weekly 2-3 part per billion (ppm) for 3 h causes repeated bouts of inflammation with progressive destruction of alveolar epithelial cells and emphysema within 6 weeks, resembling in part to that found in COPD [47,55].

I judge "ppm" is correct rather than "part per billion." Are such mechanisms active in humans at levels of 500 ppb? 100 ppb? It seems to me examination of high occupational exposures might provide some indication, but I have not found studies on occupationally exposed groups.

The Abstract of the Xu et al. (2019) paper states that the exposures to mice were single 3-hour exposures at 2.5 ppm. Changes indicating inflammation showed up with bronchoalveolar lavage (BALF). Could such changes be detected in humans as well as mice? Yes, as I learned from reading papers referenced in Appendix 3 for my response for Dr. Packham. I would like to see more evidence in the ISA of the kind in the Michaudel et al. and Xu et al papers, indicating mechanisms for health damage and biomarkers or indicators for biological changes in humans at much lower exposure levels. See my response to Dr. Packham, especially Question 3 and following. It will be important to get expert judgment on which cytokines, neutrophils and other indicators of inflammation are most significant for predicting irreversible damage to lung tissue such as described in the Michaudel et al. quote above. The support in the ISA seems weak for inflammation in humans at 60-80 ppb exposure over 6.6 hours with exercise. This support is from studies done before the last review of the ozone standard.

5. Yes, there is some useful information, and I will consider a series of published papers reporting similar findings as well-validated evidence. But the focus is not on assessing the effects of changing the NAAQS, as should be the case.
6. I will not offer an opinion on Table ES-1 beyond what I have earlier expressed about the causal framework. It seems clear that at 2.5 ppm ozone exposure causes respiratory damage in mice. NIOSH scientists have declared 5 ppm "immediately dangerous to life and health" (IDLH) for short-term exposure. But for all the other categories of causality and for ambient ozone exposure levels, I would like to see projections of what reductions in ambient levels might do to avoid adverse human health impacts, and similarly for welfare effects.

7. The overall answer is yes on all counts. But it would take me a long time to go beyond what I have already said above to describe what those changes should be. I am therefore going on to the other questions.

Responses to Questions from Dr. Mark Frampton

1. Cardiovascular (CV) effects. I have read through section ES 4.1 and Appendix 4, and selected papers included Frampton et al. (2015). I do not dispute EPA's downgrading from "likely" to "suggestive." (My answers to Dr. Cox's questions indicate my difficulty with EPA's causal determination framework.) It is my impression that the available epidemiology has many serious issues of confounding. I think these confounding issues deserve CASAC's attention, and that the ISA discussion on these confounding issues should be improved. I think research such as your own on cardiovascular effects in Frampton et al. (2015) is important, because lung damage and cardiovascular impacts are often linked in subtle ways, as indicated in Figures 4-1 and 4-6. Your 2015 paper indicates oxidative stress as a "fundamental mechanism." I think finding clear evidence of changes such as seen in rodents at 2.5 ppm to occur in humans at levels of 200, 100 ppb, and lower would be extremely important as evidence for risk of adverse CV effects in humans at ambient levels. I wish your success in your ongoing research.

2. Metabolic effects. I have a similar answer for this question. I have read through ES-4.1 and Appendix 5. I will not dispute EPA's ranking of "likely" for metabolic effects. I view potential for confounding as very high, and would like to see other variables included that are causally

related to metabolic effects. I have no additional studies to suggest. A search of the epidemiology literature will readily find many studies showing that metabolic effects depend on diet, socioeconomic status and medical care. Ozone at ambient levels might have metabolic effects, but I expect the main risk to public health to be from respiratory effects, and possibly, cardiovascular effects.

3. Total Mortality. Again, my response is similar. I am concerned about confounding, and I would like to see good biological support in clinical studies, such as your have carried out in your career, that there is significant indication of biological damage from ozone exposures in humans at levels of 100 ppb and below. There should be occupational cohorts where exposures of 100 ppb and higher with exercise occur frequently. I am concerned about the confounding with temperature, exercise patterns in indoor air and outdoor air, and air conditioning for indoor air, which reduce exposure levels considerably compared to ambient outdoor levels. Some of the highest ambient U.S. exposures occur in California, where I live. Many agricultural workers in the Central Valley of California are exposed outdoors while exercising, at levels exceeding the current MDA8 standard. I would like to see epidemiology studies carried out on this population and other populations where ozone exposure is high. (I note in Section 3.1.4.4.3, page 3-37, lines 22-23: "There are no recent studies in the U.S. or Canada that examine the relationship between short-term ozone exposure and pulmonary inflammation in healthy populations." See also 3-38, line 18; 3-47 line 6-7.) And I would like to see wildland fires viewed by EPA (and CASAC) as a major source for ozone as well as for $PM_{2.5}$. I view health risk from air pollution from wildland fires as deserving much more attention than it is receiving at the national level. The large fires California has experienced in the past three years are not natural, but the result of long-standing poor national and local policies that can be changed. It makes no sense to be spending a lot of money controlling mobile and stationary sources of ozone precursors and not spending comparable amounts of money to reduce these ozone precursors by reducing the frequency and extent of large wildland fires. These fires result from wildland fuel-buildup combined with electric power lines and other ignition sources and the dry east wind conditions that California experiences each fall. The climate aspect is getting worse, and we can expect more fires until the electric utilities do better in reducing ignitions from power lines and landowners reduce the large amounts of flammable vegetation.

Questions from Dr. Steven Packham

Question 1

I agree that there is strong evidence that ozone exposure with exercise in the range of 60 to 100 ppb causes a small but statistically significant decrease in FEV1. But whether this should be considered as an adverse health effect motivating lowering the standard from 70 ppb is a judgment call, in my opinion. Ozone exposure at levels of 5 ppm is judged immediately dangerous to life and health in humans, and animal toxicology shows clearly adverse changes at levels of 2-5 ppm. At these levels, causality seems quite clear.

Is a small change in FEV1 after 6.6 hours exposure during exercise predictive of irreversible damage? I am not a medical professional with extensive experience in treating patients, but my information from interacting with medical professionals is that small changes in FEV1 are not considered adverse if there is no lasting change – that is, FEV1 is retested and found to be within variation of earlier measurements within a few days.

A recent publication by Richard Belzer and R. Jeffrey Lewis ("The Practical Significance of Measurement Error in Pulmonary Function Testing Conducted in Research Settings," *Risk Analysis* 39(10):2316-2328) investigates the variability in FEV1 testing and protocols for summarizing data from repeated tests. I quote from the Abstract:

> Measurement error is defined as the difference from between results from the conventional protocol and an unconstrained, eight maneuver alternative. In the default model, average measurement error is shown to be about 5%. The minimum difference necessary for statistical significance is shown to be 16% … Within-day FEV1 differences ≤ 5% among normal subjects are believed to be clinically insignificant. Therefore, many differences reported to be statistically significant may be artifactual.

These statements about clinical significance supported in the main text with statements from the American Thoracic Society and the European Respiratory Society, such as at the bottom of column 2, page 2317 to 2318.
While I have read through this paper, I have not attempted to review its many references, which include the two papers Adams 2006 a and b referenced in your questions, Schelegle et al. 2009 in the ISA, and many papers on spirometry not referenced in the ISA on FEV1 or included in your questions. I urge you to read this Belzer-Lewis (2019) paper and draw your own conclusions about the significance of small FEV1 changes, especially those that are not reported as persistent over time following ozone exposure.

I hope you and others on CASAC will consider the Belzer-Lewis conclusion and recommendation:

> The failure to account for intratest variability is a material limitation of conventional spirometry in research settings. There appears to have been no systematic effort to collect sufficient data to estimate intratest variability, whether for the population, research samples, or subpopulations of interest. All spirometric protocols recognize that intratest variability is important; hence, the universal guidance to conduct multiple maneuvers. But this recognition is abandoned in practice by terminating tests early, thus failing to collect needed data, and discarding all but a single fixed value to represent each test. The result is measurement error and bias.
>
> Measurement error has pernicious effects on research intended to make causal inferences about small changes after treatment or exposure. (p. 2326.)

The authors then list recommendations for collecting further data on variability "to insure that inferences about the statistical significance of observed changes are statistically valid."

Question 2

My understanding impression is that many experts on the toxicology of PM and ozone believe PM at ambient levels poses a greater risk than ozone at ambient levels, and that the mechanisms for toxicity are

different: PM is not an oxidative challenge, and ozone is. My own view is that there is considerable uncertainty, and that treating PM as a mixture obscures the possibility that some PM exposures may be considerably more toxic than others, so that concentration response relationships in different regions with differing PM sources may vary. I commented on this for the PM draft PA.

I advocate rewriting of ES4.1 and the entire Integrated Synthesis section with much more discussion of uncertainty and variability, and clear statements of where extrapolations are being made from observations at higher doses to doses at or near ambient levels. I am disappointed that after more than a decade since the publication of the Adams and Schelegle et al. papers, CASAC does not have in the ISA better clinical information about human response to ozone while exercising than these FEV1 data, with the variability problems pointed out by Belzer and Lewis. I will discuss the ISA comments about inflammation below, as an appendix to my response to Question 3.

Question 3

I would be much more impressed by the cumulative dose argument you present with Packham Figure 1 if there were data from subsequent days indicating that significant decrements in FEV1 or in symptoms persisted. I regard FEV1 as a measurement with substantial variability, among subjects and in the responses for each subject. And I do not view transitory changes of less than 10% in FEV1 for healthy subjects as an adverse health effect motivating a national standard. After reading Belzer and Lewis I would ask whether fatigue rather than cumulative ozone dose might have contributed to larger FEV1 decrements in the latter portion of the exposure period.

I did not find Foster et al., 1987, in the ES, IS, or Appendix 4 references, but only in your questions. I note it is over 20 years old. I agree that it and similar, more recent studies in human subjects should be included in the ISA. There are at least two references to Foster's group's later work in Appendix 3, Foster et al. 1997, and Foster et al., 2000. Dr. W.M. Foster was with the Pulmonology Division at Duke University Medical Center in 2011, and his group did a number of studies on human exposure to ozone. It would seem worthwhile looking further into this group's work. Here is the web link for the 2011 paper, which is not referenced in the ISA:
https://www.physiology.org/doi/full/10.1152/japplphysiol.00337.2011.

Toward the end of the period for preparing my response I obtained a copy of the full text of the 1987 paper from Aaron Yeow. The seven healthy male human subjects were exposed to ozone levels of 200 and 400 ppb for two hours, of rest alternated with light exercise. "Mechanical and mucociliary function responses to ozone by lung airways appeared concentration dependent."

I have not been aware of the confounding benefit of 200 ppb ozone exposure in improving PM clearance. I presume this is from Foster et al. 1987. Have these Foster et al., 1987 results been replicated? I looked at titles of Foster papers subsequent to 1987 and then downloaded the "small airways " 1997 paper. Here is what I found (on p.664):

> We suggest that [the observed] response is attributable to O_3-induced alteration in bronchial tone and/or mucus secretions within smaller peripheral airways (Foster, et al., 1987,1993) similar to non-uniform ventilation associated with bronchitis and disease of the small airways. (Other references in the quote have been omitted.)

Foster contributed a book chapter, "Effects of oxidants," in *Air Pollution and the Respiratory Tract*, edited by D.L. Swift and W.M. Foster, New York: Dekker, 1999. That might be a good place to look to ascertain what Foster thought about mucociliary clearance 12 years after his 1987 publication. Without confirmation from other studies, I would hesitate to endorse that ozone exposure induces increased PM clearance, but further investigation might provide support beyond the 1987 paper that ozone exposure has potential benefit in PM clearance via increased mucus secretion.

I have read other, more recent references. Hatch et al. (2013) exposed human volunteers and rats to 400 ppb ozone with a tracer, isotope ^{18}O, and collected samples with bronchcoalveolar lavage (BALF). The authors found that:

> …resting human subjects achieve a much lower alveolar O_3 dose than exercising subjects and that this dose is comparable to that of resting rats. The resting subjects also show fewer detectable O_3-induced cellular, biochemical, and physiological (FEV_1) effects than exercising subjects.

You state on page 4, "These studies [by Folinsbee, Adams, Horstman, Kim McDonnell, Schelegle and others} document that the effect of O_3 on reduced FEV1 volumes is temporary, and suggest that hourly mean ambient O_3 concentrations below 70 ppb are not likely to cause FEV1 effects in most healthy adults." I take the "temporary" in this sentence as highly significant. I read Belzer and Lewis as indicating that much is known about FEV1 changes in sensitive subgroups. Are the changes observed in these sensitive groups from modest exposures to ozone while exercising also temporary? I am interpreting your question as motivating more balance and comprehensiveness in including toxicology, clinical human studies, and biomedical research – as well as epidemiology, in order to validate causal relationships and determine inhalation dose rates for adverse inflammatory responses in pulmonary tissues. As you might infer from my responses to the questions from Dr. Cox, I am concerned about inferences from associations in the variety of currently available epidemiological studies.

Ozone, a potent oxidizing agent, clearly seems causal for adverse inflammatory responses in human pulmonary tissue at exposure levels above 1 ppm. Exposures at or above 5 ppm are judged immediately dangerous to human life and health. But humans can remove most of low or moderate concentrations in the upper respiratory tract, except when their level of exercise requires breathing through the mouth as well as the nose. With exercise, especially vigorous exercise, the material in Hatch et al. (2013) suggests increased O_3 content in BALF cells at 400 ppm might indicate adverse effects such as inflammation. What evidence other than the temporary FEV1 changes do we have for adverse effects in the 60 -100 ppb range? It seems to me that inflammation at near-ambient ozone levels is an obvious place for CASAC to focus its attention and ask EPA for further information and evaluation.

Support for an important statement in the IS Overall Conclusions Table, first bullet, second sentence:

> "The strongest evidence comes from controlled human exposure studies demonstrating ozone-induced decreases in lung function and inflammation in healthy, exercising adults at concentrations as low as 60 ppb after 6.6 hours of exposure."

Has inflammation been demonstrated in healthy exercising adults at concentrations as low as 60 ppb, as stated in this Table, page IS-1? Affirmative answers are in the text at IS-7, lines 5-6, and also 3.1.4.5, page 3-38, lines 27-29 but without listing the exposure level. Similarly in IS 4.3.1, page IS-23, current evidence is summarized as follows:

> There are, however, no new 6.6-hour ozone exposure studies since the 2013 Ozone ISA. Evidence in the 2013 Ozone ISA demonstrated increases in FEV1 decrements, respiratory symptoms, and inflammation following ozone exposures of 6.6 hours, with exercise, as low as 60 to 70 ppb (Section 3.1.4).

A similar statement, lacking the qualification that exposures were with exercise, is found at 3-80, lines 29-32:

> Controlled human exposure studies demonstrate ozone-induced decreases in FEV1 and pulmonary inflammation at concentrations as low as 60 ppb after 6.6 hours of exposure. The combination of lung function decrements and respiratory symptoms has been observed following 70 ppb and greater ozone concentrations following 6.6 hour exposures.

What is in 3.1.4 to support this statement?

Controlled Human studies are discussed in 3.1.4.4.1. Here is the first paragraph:

> As reported in studies reviewed in the 1996 and 2006 ozone AQCDs (U.S. EPA, 2006, 1996a), acute ozone exposure initiates an acute inflammatory response throughout the respiratory tract that has been observed to persist for at least 18−24 hours post-exposure. A single acute exposure (1−4 hours) of humans to moderate concentrations of ozone (200−600 ppb) while exercising at moderate to heavy intensities results in a number of cellular and biochemical changes in the lung, including an inflammatory response characterized by increased numbers of PMNs, increased permeability of the epithelial lining of the respiratory tract, cell damage, and production of proinflammatory cytokines and prostaglandins. These changes also occur in humans exposed to 80 and 200 ppb ozone for 6−8 hours.

This paragraph begins by citing EPA documents from 1996 and 2006, based on studies up to that time. Let me focus on the last sentence: “These changes also occur in humans exposed to 80 to 200 ppb ozone for 6-8 hours.” Is the support at 80 ppb from old studies or new ones? The following paragraph mentions 200 ppb, but not 80 ppb. The study source of the increases in sputum polymorphic neutrophils (PMNs) in EPA (2013a) is not given. Then the section text has bullet points about specific newer studies.

The first bullet discusses PMNs and shed epithelial cells observed in BALF. Alexis et al. (2013) would appear to be the source for the statement about increased sputum PMNs at 60 ppb. Then comes Arjomandi et al., (2018), the only reference dated after 2013-4. It is a larger study of healthy older adults. Here is what the authors of that paper write about their findings at 70 and 120 ppb exposures in older adults:

> Ozone exposure caused a marginally significant increase in PMN% in a concentration-dependent manner (P = 0.012) (Tables 2 and 3, and Figure 3). In the mixed model regression, PMN% increased by an absolute value of 8.16% after 120 ppb compared with 0 ppb ozone (95% CI, 2.84–13.48%; P = 0.003) (Table 3). The 70 ppb effect was not significantly different from 0 ppb (P = 0.134). The absolute PMN count (ln PMN/mg) also showed a positive but nonsignificant increase with increasing ozone concentration (Table 3).

I interpret these statements as indicating a "marginally significant" response at 120 ppb and not supporting a clear finding of an inflammation response at 70 or 60 ppb. However, there is a s statement at 3-30, lines 22-25 that the delivered dose of ozone in the Arjomandi et al study at 120 ppb might have been only 60% of the delivered dose of ozone at 60 ppb in the Kim et al. study. In the Arjomandi et al. study a plasma CC!6 response after exposure at 70 ppb with exercise was not significantly different from control at 0 ppb.

In the second and third bullet, Bosson et al. (2013) measured blood neutrophils for up to 18 hours after ozone exposure, at 200 ppb for 2 hours, compared to filtered air. There was a transient decrease in blood neutrophils. The authors write:

> To date, although numerous groups have examined ozone-induced systemic inflammation in humans by measuring inflammatory mediators and PMN priming, no simple description of the changes in cellularities has been published. … At 18 hours post exposure we found no evidence of blood neutrophilia, though inflammation persisted in the airway lumen at this late time point. No other changes in peripheral blood cell types, or lymphocyte subsets were noted at any of the measured time points." … "We therefore believe that there is merit in further exploring the relationship between systemic neutrophilia and ozone in the real-world setting."

These statements do not indicate strong evidence of inflammation at 70 or 60 ppb, but rather that what is known now motivates further research. It is consistent with a preceding statement in the ISA, at the end of 3.1.4.2.1, page 3-23, line 22-26:

> Although several studies have investigated the effects of 6.6-hour exposures during moderate exercise to 60 ppb ozone, none have observed a statistically significant increase in respiratory symptoms following ozone relative to filtered air. There are no new controlled human exposure studies conflicting with the above or contributing a better characterization of ozone-induced respiratory symptoms.

But on 3.1.4.2.4. page 3-24, lines 18-22, the ISA claims support for an "adverse response," based on a model from 2013:

> A recent model can be used to determine the ozone concentration that would lead to the same FEV1 decrement following an 8-hour exposure (McDonnell et al., 2013). Under the assumption that respiratory symptoms might accompany similar ozone-induced FEV1 decrements, regardless of exposure duration, the model indicates that an 8-hour exposure to 64 ppb ozone concentration might reasonably be expected to cause an adverse response in young healthy adults.

I read McDonnell et al 2013. The model is of about intrasubject variability in FEV1. In the Appendix the authors state:

> One could interpret X to represent the level of oxidant stress resulting from accumulation and removal of ozone or its reactive byproducts, although the validity of the model is not dependent upon this interpretation.

But all the data used is FEV1 data, and X is a dose rate for ozone, concentration times minute ventilation with adjustments for body surface area, age, and body mass index. There is no tie to any measure of inflammation or to respiratory symptoms. The "interpretation" of oxidant stress is in no way confirmed by any data. And I did not find the phrase "adverse response" defined in this paper, which is about modeling between-subject FEV1 response variability. "Adverse response" is not defined in the ISA either, as far as I can determine. There is no support I could find in the McDonnell et al. (2013) paper for the "assumption that respiratory symptoms might accompany similar ozone induced FEV1 decrements."

I return to 3.1.4.4, the fourth bullet point about inflammation and oxidative stress, page 3-29. Here the ISA notes the studies by Fry, et al. (2012) and (2014) and Hernandez (2012) "included no filtered air control arm," and "without an air control it is not possible to assess potential effects of exercise and/or the laboratory procedures on results." The ISA notes (p. 3-29, lines 42-44) regarding comparison of lung function with inflammation indicators, "This is consistent with studies reviewed in the 2006 Ozone AQCD showing spirometric measures and inflammatory responses to ozone are unrelated." (This would seem to contradict the "assumption" from MacDonnell et al. (2013).)

In six years since 2013 the information in the ISA indicates there have not been clinical studies that have confirmed the "predictions" of the 2013 model developed by McConnell et al., or confirmed the findings of inflammation biomarkers with exposures of 60 ppb for adults exercising over a 6.6 hour period from Kim et al. (2011) and Alexis et al (2013). For the second sentence of the "Overall Conclusions" of the ISA, page IS-1, the support seems weak. And this evidentiary support is not new, but was available at the time of the last ozone standard review.
Dr. Packham, you are an expert for pulmonary physiology and inhalation toxicology, and so your understanding of the available published studies should be much better than mine. I hope you will review this material in Appendix 3 plus other relevant clinical studies in detail. These studies seem of considerable importance in judging whether the current MDA8 ozone standard at 70 ppb is adequately protective of human health.

Questions from Dr. Sabine Lange

1. Statistical significance is useful in epidemiological studies, but in a limited way. These studies use regression to determine the association between a predictor x and a consequence y. This association might be statistically significant, that is, a good predictor, but causality could be absent: There may be a cause z that affects both x and y. Two examples: Children's shoe side predicts the children's reading ability. (Example due to Judea Pearl, in *The Book of Why*). Ice cream sales predict heat stroke cases. (I think Dr. Cox uses this one, which was more accurate before the era of air conditioning.) Progress for better prediction is to consider that there may

be other factors that are predictive of y, get the data on these, and use these data on making the prediction. Children's age and high ambient temperature are candidates for the two examples.

2. (and 3). I'm not a fan of these cross-over studies, especially for mortality as the end point. For a good example, consider exposure to wildfires such as we have been experiencing in California. Consider hospital admissions for respiratory distress. What was the level the day before the smoke plume affected the area? What was the level the day the plume arrived? And after the wind blew the smoke away, then what was the level the next day? Yes, one might expect a low response level before the plume, and high levels after the high exposures, perhaps persisting for days after the levels have dropped. For Q3, the response is number of deaths on the exposed day versus the control day, and not the death of an individual person. (With cohort studies, it is more complex.)
3. I read the Di et al. study. Co-authors Schwartz and Zanobetti are among those trying to figure out how to do epidemiology where additional factors are considered. But I am not persuaded that confounding was not a significant issue for the results in the Di et al. study. The data base was all Medicare patients who died in a twelve year period. Most of the deaths occurred on days with ozone and $PM_{2.5}$ levels well below the current standards. The death rates per 10 µg/m3 for $PM_{2.5}$ was 1.45 per million persons at risk per day, and for 10 ppb ozone, 0.66 per million. These are extremely small numbers, but with sample size of nearly a hundred million days, the confidence limits were narrow around these numbers and did not include no increased risk. I looked up reference 9, Maclure (1991), on the study design. The Maclure Abstract begins, "A case-control design involving only cases may be used when brief exposure causes a transient change in risk of a rare acute-onset disease." I don't see the biological plausibility of comparing case days and control days for total mortality – not a rare acute–onset disease, but rather a situation where people who may be already very sick tend to die on days when they have additional stress. I suspect that high temperature may have acted as a confounding variable. Looking at figure 5 in the Di paper, I notice that the exposure response curve seems to flatten out (i.e., is insensitive to exposure level) for the higher 50% of the exposures, both for ozone and for $PM_{2.5}$. If these pollutants were causing the mortality increase, I would expect that the lower half of the exposure levels would be the flatter portion, and at higher exposures there would be more of a positive concentration response relationship. What may be going on is that in the days with higher half of the exposure levels, the pollutant levels are correlated (but rather poorly) with the frequency of very hot days. On such very hot days mortality is significantly elevated. But on the lower pollution half of the days, there is a stronger correlation: a much lower frequency of very hot days. Very hot days can cause stress to an elderly person, especially in non-air conditioned space. Remember, the Di et al response rates are on the order of a one-in-a-million change. A small number of very hot days correlated with elevated exposure levels might give results such as reported in this paper.
4. I view it as extremely important that ozone concentrations are 2 to 10 (or more) times less indoors, and where the indoor space is air conditioned this has a big influence on 2 versus 10. Who is then at risk with high outdoor exposure during exercise? In California's Central Valley, with high peak ozone exposures, it will be agricultural workers, especially for hand harvesting of crops. In urban areas, we might expect that high exposures will occur for hikers, runners, bicycle riders, garden workers, and children playing outdoors.
5. That is a good question. My impression is that at ppm levels, humans and rodents are about equally sensitive, and that prolonged exposures at 5 ppm or higher are life-threatening to

humans. There seems to be some information at lower levels. See my responses to Drs. Cox and Packham.

6. Causality diagrams are still rare. But some epidemiology studies do consider multiple predictive factors, and explain how they do it. I expect we will consider this aspect in the risk assessment in the upcoming Ozone PA.

Questions from Dr. James Boylan

Appendix 1: This is not my area. I would like to see more on wildland fires as an important ozone source. There are references to good papers by Daniel Jaffe. I would like more on the impact of strategies to reduce ozone from fires (wildlland and other) and how background ozone levels combine with ozone from anthropogenic sources in the US. Changing the NAAQS does not change background. Calculating changes in ozone exposure under a revised standard needs to be done, and should be done with realistic rather than simplistic assumptions. Such calculations should be presented in the upcoming Ozone PA.

Appendix 2. Again, it's not my area. Due to time limitations I only skimmed this appendix. It is my impression from reading it that there is an abundance of detail on geographical variation, and not enough emphasis on uncertainty and variability in the patterns of indoor versus outdoor exposure in relation to ambient monitoring, and also exercise patterns, for connecting ozone exposure (especially as measured in epidemiology studies) with human health effects.

Appendix 9. I read Appendix 9 and found it to be generally well done. I chaired the EPA Science Advisory Board's review of the two reports by EPA to Congress on global climate alteration in 1989-1990. I have no specific suggestions regarding accuracy or completeness. While I have followed the extensive literature on climate alteration over the past two decades, I am not familiar with the far smaller number of publications specially addressing tropospheric ozone for radiative forcing, climate alteration, and impacts of UV-B shielding on human health and ecosystems. I believe it is useful to have this chapter included, as I believe it covers important topics for EPA under the provisions of the Clean Air Act.

Questions from Dr. Corey Masuca

1.3.1. It is my impression that volatile organic compounds (VOC) and nitrogen oxides (NOx) are the main precursors of ozone, but that CO and CH4 also contribute. I am not expert in this chemistry. CH4 is an important gas for climate change/greenhouse warming. I do not believe CO is considered to be so, in part because it is removed from the atmosphere much more quickly than CH4. There is discussion in Appendix 9 on the role of CO and CH4 for radiative forcing and climate alteration that you might wish to read. See page 9-9 line 2 through 9-12 and Figure 9-3, 9-4 on pages 9-11 and 12.

1.3.1.2.1. Repeating my answer above, methane is a minor contributor when VOCs are plentiful, but contributes significantly when VOC levels are low. See 1-23, lines 27-30 and Appendix 9. NOx formed by lightning reacts with CH4 in the atmosphere to make ozone, shortening the atmospheric lifetime of the CH4.

1.3.1.2.2. Transport from Asia is significant for ozone in California and other western states. See the background ("USB") discussion in this section and for more detail, the Jaffe et al. (2018) paper. Regional transport of precursors is important. There is a lot of literature on how to model ozone formation based on precursor emissions. Such modeling is complex, and is not an area in which I have much expertise.

1.3.1.3.2. Biogenic VOC such as terpinoid compounds can be important for ozone formation, especially in the southeast U.S. The Great Smoky Mountain National Park is a classic example, because the name comes from smog formation from the biogenic VOCs. Here is a link to an article you might enjoy reading: https://www.livescience.com/46958-trees-ozone-pollution-map.html. More generally, biogenic VOCs can be the dominant contributors to ozone formation outside urban areas. See page 1-20 lines 9-15. In many urban and suburban areas of the US these biogenic VOCs are less important than the local anthropogenic sources of VOC. One must consider both the VOC and NOx levels, as either can be limiting in ozone formation chemistry. See, for example, the discussion on page 1-52, lines 13-18. In the past highway vehicles were main sources of NOx and VOC, but these emissions have been greatly reduced, as shown in Figure 1-3, page 1-11.

1.4. and continuing on 1.3.1.3.2. Situations occur where ozone forms downwind of an urban area when an air rich in NOx reaches these biogenic VOCs: There was not enough VOC over the urban area to enable the ozone formation to occur earlier. The chemistry and air flow movements from meteorology make the modeling of ozone formation quite complicated. Humidity is one of many important factors. I am not knowledgeable on how well modeling predictions are confirmed via remote sensing and local monitoring. I am doubtful that specialized monitoring of individual chemicals would be very helpful in improving prediction of ozone levels, but my knowledge in this area is quite limited. The ISA at 1-21, lines 21-27 agrees with me for biogenic VOCs. For some anthropogenic source precursor chemicals the ability to predict may be much better.

1.5 Yes, inversion layers are important. We in California have a "mountain bowl" around Los Angeles, and a much bigger one in our Central Valley, where many the highest MDA8 ozone levels in the nation occur.

The discussion of these issues in the ISA Appendix 1 is highly detailed. Based on my limited knowledge the sections you refer to in your questions above seem generally accurate.

Exposure monitoring for ambient and indoor ozone is not an area where I have expertise, and so I am skipping most of your question on Appendix 2. In my judgment this appendix is overloaded with detail, instead of emphasizing the large differences between indoor ozone exposure with and without air conditioning, and ozone exposure in outdoor ambient air. Indoor exposure may be comparable to outdoor exposure on cool days with open windows. On warmer, higher outdoor ozone days, the indoor levels may be ½ to about 1/20 the outdoor levels, with the lower levels in indoor space with air conditioning and filtration. These big differences for human exposure are not apparent in the text and tables. (The numbers I cite above are roughly consistent with Table 2-4.)

Regarding 2.4.1. Many people will prefer to avoid outdoor exposure and stay indoors in air conditioned space on extremely warm/hot/humid days. For some people, exercise outdoors even in extreme weather is essential. And for many lower income people, including children and elderly people, their homes,

workplaces, and schools may not be air conditioned. Heat stress and lack of adequate hydration can bring on additional mortality and morbidity, independent of ozone exposure. I view this potential for confounding as a very important issue for interpreting the epidemiological studies. As one example, Paris, France is a modern city and not known for high levels of air pollution. The heat waves this summer (2019) have reportedly caused nearly 1500 deaths. A heat wave in Paris in 2003 killed about ten times as many (15,000): https://www.bbc.com/news/world-europe-49628275. We should recognize that persistent, stagnant hot air – a "heat wave" - can cause fatalities, independent of the effects of air pollution. And as you point out, such stagnant hot air days are often times of high ozone concentration levels.

Miscellaneous Questions(s)

Yes, continuing from the response above, I think assessing impacts on lower socioeconomic status (SES) populations is very important. Mortality is elevated in lower SES populations. Is this elevation from higher ozone exposure? Or is it from less medical care, poorer housing, and lack of air conditioned space at home and at work? These interacting factors should be examined so that confounding is minimized.

We need to help our lower SES populations, such as those people with lower incomes and their children. We do not want high numbers of deaths in heat waves, such as have occurred in Paris. French authorities claim they were better prepared in 2019 than in 2003 – yet a lot of people still died this year.

References not in the ISA or Questions from CASAC Members:

Richard B. Belzer and R. Jeffrey Lewis (2019), "The Practical Significance of Measurement Error in Pulmonary Function Testing Conducted in Research Settings," *Risk Analysis* **39**(10):2316-2328.

Questions from Dr. James Boylan

Appendix 1 – Atmospheric Source, Chemistry, Meteorology, Trends, and Background

• *Is the discussion on metrics and definitions (Section 1.2) accurate and complete? If not, what additional information needs to be included?*

Section 1.2 is complete and accurate with a couple of exceptions.

Section 1.2.2.1 states that "… USB is a model construct that cannot be measured using ambient monitoring data." This statement is correct, in that USB cannot be **directly** measured; however it can be estimated from ambient monitoring data as shown by Parrish et al. (2017) and Parrish and Ennis (2019). This will be discussed in more detail below in response to the question regarding Section 1.8.

Section 1.2.2.5 discusses baseline ozone, which is defined as "the measured ozone concentration at rural or remote sites that have not been influenced by recent, local emissions (Jaffe et al., 2018)." This definition would be more accurate if "local emissions" were replaced with "local continental effects". Ambient ozone concentrations are affected not only by emissions, but also by surface deposition, particularly to vegetation. Measured concentrations at rural or remote sites can be greatly reduced from true baseline concentrations by surface deposition even in the absence of emissions. This is an important issue.

• *Is the discussion on sources of U.S. ozone and its precursors (Section 1.3) accurate and complete? If not, what additional information needs to be included?*

Generally this entire Section 1.3 has an important shortcoming. On pg. lxxvii of the Preface the purpose of the ozone ISA is stated: "This ISA communicates critical science judgments of the health and welfare criteria for ozone, and serves as the scientific foundation for the review of the current primary (health-based) and secondary (welfare-based) National Ambient Air Quality Standards (NAAQS) for ozone." In reading through Section 1.3 a great deal of scientific information is summarized, but there is little or no discussion of the relevance of this science to the NAAQS or the ozone design values upon which the NAAQS is based. For example, wildfires (here included in the broader category of landscape fires) and stratosphere-troposphere exchange are discussed. These are two natural sources of ozone that are specifically addressed in EPA's exceptional events rule. A crucial issue is the extent to which these sources can affect the ozone design value and perhaps cause an exceedance of the NAAQS that would be eligible for addressing under that exceptional events rule. How often do either of these natural sources cause exceedances? It is difficult to evaluate the significance of a particular scientific issue without the context of how that issue might affect the NAAQS that is being reviewed.

Another significant shortcoming in Section 1.3 is that the uncertainty of the ozone precursor emissions estimates should be clearly discussed and defined to the extent possible. The words "approximately" and "estimate" are often used, but such terms are not defined. I think that a few paragraphs should be devoted to a discussion of emission inventory uncertainty; this discussion might be based on material in

Miller et al. (2006) or a similar, more recent emission inventory assessment. (This comment repeats a comment that I made in my response to a similar question regarding the PM PA; more details are given there.)

Section 1.3 discusses the role that global methane plays in the global tropospheric ozone budget, which is based on chemistry-climate model simulations. Simulations by different models generally agree, but the models generally use similar parameterizations of critical physical processes. In my opinion, increasing methane may increase global ozone concentrations, but due to model uncertainties that relationship is uncertain. Conceivably increasing methane may decrease, rather than increase, global ozone concentrations. (See expanded discussion in my response to a similar question from Dr. Corey Masuca.)

Section 1.3.1.3.1 mentions that "soil NO emissions rates can be high enough to affect local and regional ozone concentrations under certain circumstances", but largely dismisses the importance by noting that "Biogenic emissions of NO_X are estimated to contribute only a small part to national NO_X emissions, …." However, there is evidence (Oikawa et al., 2015; Almaraz et al., 2018) that NOx emissions from agricultural soils plays a major role in degraded air quality in the Imperial Valley, located in an ozone no-attainment area in California. This issue deserves more discussion since it may play a significant role in an important no-attainment areas in the country.

Section 1.3.1.3.3 Landscape Fires deserves improvement. It provides a synopsis of literature results, but there is no synthesis of the current state of knowledge. A summarizing statement in the context of recorded ozone design values would be valuable. It is stated that "*Wildfires contribute a few parts per billion (ppb) to seasonal mean ozone values in the U.S., but episodic contributions may be as high as 30 ppb*." My understanding is that wildfires vary in location and intensity from year to year, but have been generally increasing over the last decade or two. However, in one example northern rural state where high impacts of wildfires are expected (Montana), the ozone design values at all monitors in the state have averaged 55.4 ppb with a standard deviation of 2.2 ppb and no apparent trend over the 1979–2017 period, i.e. over nearly 4 decades (Parrish and Ennis, 2019). Similarly small variability in ozone design values without an apparent trend are found in the Dakotas. If wildfires have relatively large, episodic effects, why are these effects not seen in the ozone design values in these rural northern U.S. states where wildfire impacts are expected to be most obvious?

Section 1.3.2 discusses the influence of stratosphere-troposphere exchange processes. In reading this section one gets the impression that ozone from the stratosphere is an episodic phenomenon that has only occasional impacts on surface ozone. However, ozone from the stratosphere enters the upper troposphere where the lifetime of ozone is weeks. Hence, even if a deep stratospheric intrusion does not reach the surface, it does mix into the background troposphere, and that contributes to U.S. Background Ozone (USB), even if the stratospheric source of that ozone is not obvious. Hence the influence of stratospheric ozone is ubiquitous (often 10's of ppb), but highly variable in time. This section should make that clear.

• *Is the discussion on ozone photochemistry (Section 1.4) accurate and complete? If not, what additional information needs to be included?*

Page 1-25, line 32 refers to "NO_X as an oxidant". That is incorrect; hydroxide and peroxy radicals are generally the oxidants in ozone photochemistry. The role of NO_X is better characterized as a catalyst.

The final paragraph of this section is entitled: "*Oil and gas sector impacts on ambient ozone levels extend beyond wintertime ozone episodes."* This section cites two studies suggesting that "on specific summer days oil- and gas-related precursor emissions could contribute locally up to 30 ppb ozone." However, efforts to find evidence for increases in ozone design values that can be attributed to oil and gas development have failed. For example, ozone design values in North Dakota have remained nearly constant across the state (59.3 ppb with a standard deviation of 2.7 ppb) with no significant trend over the 1982–2017 period (Parrish and Ennis, 2019). During this period the Bakken oil and gas field was developed and began producing, with no discernable effects on ozone design values, either at monitoring sites located within the oil and gas development region, or in the surrounding areas. Similar results were found for Texas oil and gas fields.

• *Is the discussion on inter-annual variability and longer term trends in meteorological effects on anthropogenic and U.S. background ozone (Section 1.5) accurate and complete? If not, what additional information needs to be included?*

The discussion in Section 1.5 is generally accurate. What is missing is a clear, quantitative summary of the impact that inter-annual variability has on ozone design values, which are the most relevant statistic from a NAAQS perspective. Section 1.5.1 discusses meteorological effects on ozone concentrations at ground level by citing three studies that investigate mean max 8-hour ozone concentrations in various parts of the country. Section 1.5.2 discusses inter-annual and multidecadal climate variability in a qualitative manner, with only brief mentions of quantitative impacts on maximum daily 8-hour ozone.

Parrish et al. (2017) Parrish and Ennis (2019) show that the a very large fraction of the variance in ozone design values recorded in the largest U.S. urban areas over the past two to four decades is dominated by changes in precursor emissions, and that interannual variability plays only a small role. For example in six of California's southern air basin (including Los Angeles's South Coast Air Basin) simple exponential fits based on decreasing U.S. anthropogenic emissions of ozone precursors account for 98.4% of the variability of ozone design values in that region (Parrish et al., 2017). In the three rural northern states discussed earlier (Montana and the Dakotas) ozone design values over 3 to 4 decades have remained constant within standard deviations of 2 to 4 ppb (Parrish and Ennis, 2019). In the northeastern U.S. similar simple exponential fits accurately capture the temporal variation of ozone design values over 18 years (2000-2017) with root-mean-square deviations of only 2 to 5 ppb. Clearly inter-annual variability and longer term trends in meteorological effects do not contribute large variability to ozone design values. This should be made clear in the discussion.

• *Is the discussion on measurements and modeling (Section 1.6) accurate and complete? If not, what additional information needs to be included?*

With regard to measurements, ground-based ozone lidar instruments are providing a new dimension to our understanding of tropospheric ozone, as these instruments can map out the vertical structure of ozone, and quantify the mixing of plumes aloft to the ground, where the ozone they carry impact the population and surface measurements (e.g., Langford et al., 2017). A review of these instruments and their capability should be added to this section.

With regard to modeling, an overall synthesis of the current state of model performance is lacking. For the development of state implementation plans (SIPs), EPA requires regional chemical transport modeling to demonstrate that proposed emission reductions will lower the ozone design values within the state to the extent that the NAAQS will be achieved. A clear quantification of the accuracy that can be expected from these models must be included. Currently, a summary of model improvements and some model intercomparison studies are discussed. This discussion includes statements such as:

> "The normalized mean error for hourly ozone ranged between 21 and 47 ppb, while the normalized mean error for the 1-hour max ozone concentration ranged between 19 and 22 ppb (25th−75th percentile of reported studies)."

When attainment vs. nonattainment decisions are based on differences of only 1 or a few ppb, these apparently large errors in modeling would seem to indicate that the current generation of regional chemical transport models are inadequate for the task of developing realistic SIP plans. In my view, a much more comprehensive discussion of modeling is required, and that discussion should give a clear, concise summary of the accuracy of model results, shortcomings of the models, and how these shortcomings are likely to affect SIP modeling.

• *Is the discussion on ambient air concentrations and trends (Section 1.7) accurate and complete? If not, what additional information needs to be included?*

The discussion is accurate, but a tremendous resource remains unexplored. With over 1000 monitors collecting ozone data for at least the warm-season months throughout the country over many years, this data set has tremendous potential for informing us about the sources and sinks of ambient ozone in the U.S., and how these vary between regions of the country, land use, ecological environment, and many other variables. Research into this data set is important if the country is to develop ozone control policies that are protective of public health and ecological health, and at the same time economically efficient and feasible to implement. Some recent analyses based on this data set is discussed in response to the next question regarding U.S. background ozone concentrations.

• *Is the discussion on U.S. background ozone concentrations (Section 1.8) accurate and complete? If not, what additional information needs to be included?*

In my opinion the discussion of U.S. background ozone concentrations (Section 1.8) is neither accurate nor complete. The entire discussion should be written from a different perspective. Two or three decades ago urban ozone concentrations were high enough that on days with the highest ozone, background ozone made a relatively minor contribution. However, emission controls have reduced ozone in the U.S. to the extent that now background ozone contributes the majority of urban ozone concentrations, even on most days when ozone exceeds the NAAQS.

The following 2 figures show estimates of the ozone design value (ODV) that would be present in the absence of U.S. or North American anthropogenic emissions.

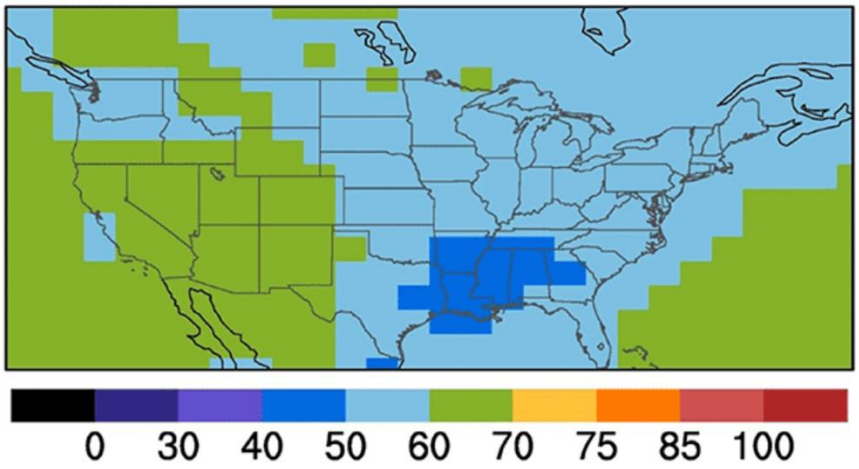


Figure 1. Annual 4th highest MDA8 O_3 in ppb from North American background (i.e., with North American anthropogenic precursor emissions set to zero) averaged over 2010–2014 from a GFDL-AM3 model simulation (Jaffe et al., 2018).

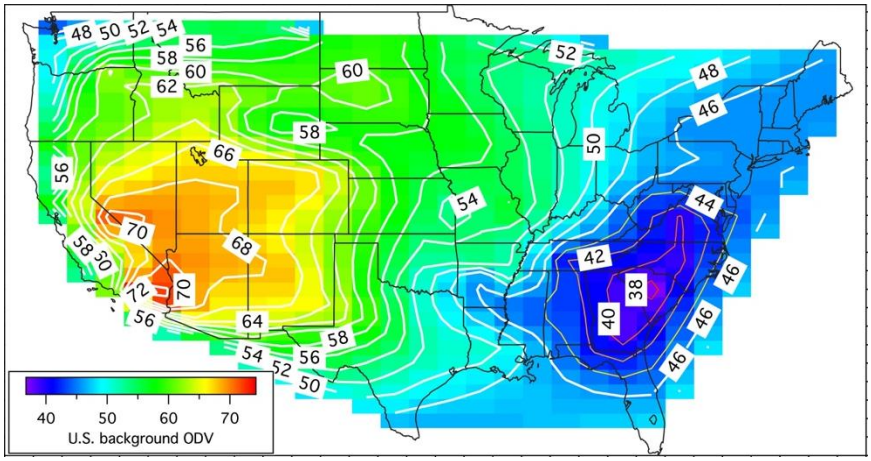


Figure 2. Ozone design values expected from U.S. background (i.e., with U.S. anthropogenic precursor emissions set to zero) in ~ 2015 derived from observations (D.D. Parrish, unpublished figure).

Figure 1 is from a model calculation using the "zero-out sensitivity approach" as described in Section 1.8.1.1. This figure is a summary plot included in a recent assessment of background ozone over the U.S. (Jaffe et al., 2018). Figure 2 is developed from an observational-based approach (Parrish et al., 2017; Parrish and Ennis, 2019) applied to the entire country. Both of these maps have unquantified uncertainties, but the agreement between them regarding the general features and their magnitude is encouraging. A critical message from these two maps is that in the southwestern U.S., background ozone makes such a large contribution (ODV up to ~70 ppb) that it will be extremely difficult (perhaps impossible) to reach the 70 ppb NAAQS, unless the background contribution decreases. This background contribution is largely unaffected by U.S. precursor emission controls. The much lower background ozone concentrations in the eastern U.S. makes it easier to achieve the NAAQS in that part of the country.

The high estimated ODVs shown in Figs. 1 and 2 for the southwestern U.S. can be evaluated by examining ODVs recorded at CASTNET (Clean Air Status and Trends Network) sites. These sites are intended for measuring rural, regionally representative O_3 concentrations and assessing changes in

background O_3 contributions. The map in Fig. 3 shows the locations of five CASNET sites in the southwestern U.S. selected to 1) be as isolated as possible from urban areas, 2) have relatively complete ozone measurement records, and 3) lie at similar elevations. The ODVs recorded at these five sites (Fig. 3) are similar among all sites (overall average 70.0 ppb with a standard deviation of 3.0 ppb) and follow a very similar long-term change (increasing before the mid-2000's, reaching a maximum, and decreasing after that maximum). Importantly, this long-term change is very different from that of the U.S. anthropogenic ozone contribution, which has continually decreased over this same period. The similarity in the ODVs and their long-term changes at these sites, which are separated by several 100s of miles, indicate that the ozone concentrations measured at these sites are approximately direct measures of the U.S. background ozone concentration. Hence, these measured ODVs support the estimates given in Figs. 2 and 3. It is significant that from ~1995 to ~ 2010 the ODVs recorded at these relatively isolated rural sites would all have frequently exceeded the 70 ppb NAAQS, solely from U.S. background ozone.

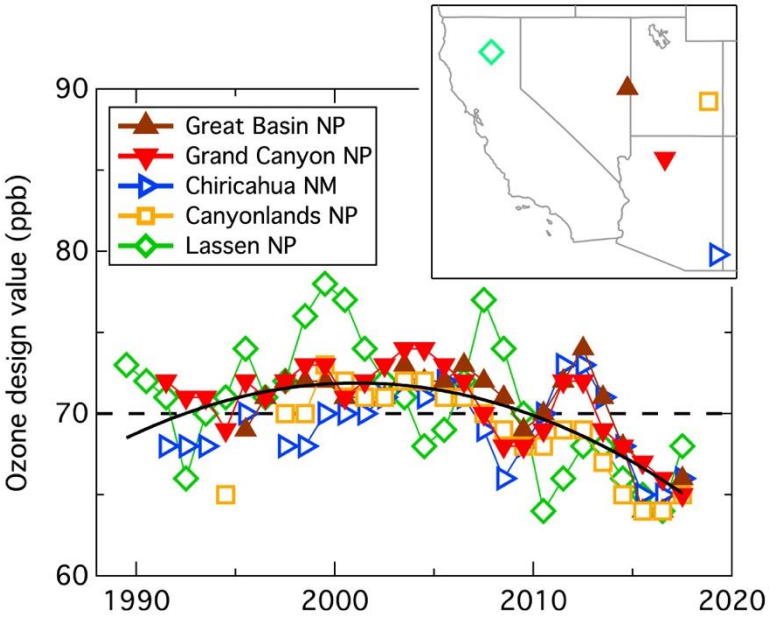


Figure 3. Ozone design values recorded at five relatively isolated CASTNET sites in the southwestern U.S. (Data from EPA's AQS data archive (https://www.epa.gov/aqs).

Figures 1 and 2 show that U.S. background ozone concentrations differ markedly between Los Angeles in southern California and New York City in the northeastern U.S. The observational-based approach of Figure 2 estimates that in the absence of U.S. anthropogenic ozone contributions, the ODV in Los Angeles and New York City would be 62 and 46 ppb, respectively (Parrish et al., 2017; Parrish and Ennis, 2019). Figure 4 shows the difference this has made in each city's efforts to attain the NAQS. In each of these metropolitan areas, the ODVs have decreased at similar rates after ~2000, with the ODVs in New York City projected to decrease to 70 ppb in the early 2020's. However, it is much more difficult for Los Angeles to attain the NAAQS because the U.S. background ozone contribution is much larger, only 8 ppb below the NAAQS. Consequently, it is projected that the maximum ODVs in the Los Angeles urban area will not attain the NAAQS until ~2050 (assuming that the U.S. background contribution remains constant and the current trend of reduction of the US anthropogenic contribution can be maintained through continually strengthened emission controls).

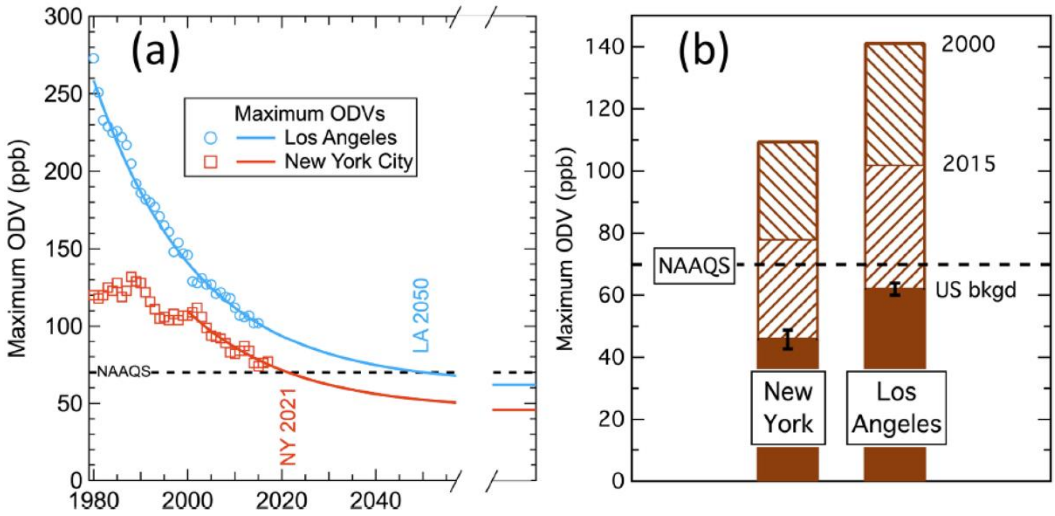


Figure 4. Comparison of maximum ozone design values (ODV) observed in the Los Angeles and New York City urban areas. (a) Temporal trend and parametric fit; annotations indicate year that extrapolations decrease to 70 ppb. (b) Bar graph indicating maximum ODVs in 2000 and 2015 (hatched bars) and the estimated US background (bkgd) ODV (solid bars) (Parrish and Ennis, 2019).

In addition to failing to clearly include the perspective described in the preceding paragraphs, Section 1.8 has other shortcomings. Section 1.8.1 begins with the statement "As described in Section 1.2.2.1, USB ozone cannot be reliably estimated using ambient monitoring data because monitors can be influenced by U.S. emissions, including both relatively nearby emissions and interstate and hemispheric transport of ozone produced from U.S. emissions." As Parrish et al. (2017) and Parrish and Ennis (2019) show, and as illustrated in the 4 figures above, USB ozone can indeed be reliably estimated using ambient monitoring data. Monitors can be (and usually are) influenced by U.S. emissions, but it is possible to account for these influences. The estimates from this measurement-based approach and from the modeling-based approaches have significant uncertainties; our best avenue to obtaining reliable estimates of U.S. background ozone with minimum uncertainty is to use both approaches, and to attempt to understand differences in the estimates when differences arise. Section 1.8 simply ignores the measurement-based approach based on insufficiently supported reasoning (Section 1.8.1.4).

Section 1.8.1.5 discusses uncertainties in and disagreement among model results. A useful summary statement is provided: "… USB ozone estimates contain uncertainties of about 10 ppb for seasonal average concentrations, with higher uncertainty for MDA8 average concentrations." Modeling MDA8 concentrations on the specific 4 days per year with the largest background ozone are required for determination of the concentrations illustrated in Figure1; thus the uncertainties of such determinations from models must be larger than 10 ppb. Since attainment vs. non-attainment decisions are often based on only 1 or a few ppb, the usefulness of these model estimates is greatly limited by this uncertainty. Section 1.8.1.5 should emphasize that this state-of-the-science of the models is a serious impediment to our understanding of U.S. ozone concentrations.

Section 1.8.2.1 discusses new USB and North American Background estimates, but all of these estimates are for seasonal means. It is critical to evaluate the ODV that can result from USB, because

that informs us of the "headroom" available for anthropogenic ozone production on those high background days. This concept is illustrated in Figure 4 above.

Section 1.8.2.3 discusses the contribution of USB to ambient air ozone as a function of ozone concentration. The discussion misses an essential point: as the ODV in an area is reduced and approaches the NAAQS, the relative importance of the USB will increase, because the USB contribution will remain (approximately) constant while the anthropogenic contribution decreases. For modeling relevant for reducing the ODV to concentrations equal to or below the NAAQS, one should focus not only on today's high ozone days, but also on the days in the future that determine the ODV when it is still slightly above the NAAQS. It is likely that those days will have a much higher relative contribution from USB, much as the situation in the high-elevation locations in the western U.S described in this section; at those locations USB is consistently predicted to increase with total ozone concentration.

Appendix 2 – Exposure to Ambient Ozone

• *Is the discussion on exposure concepts (Section 2.2) accurate and complete? If not, what additional information needs to be included?*

I have no relevant expertise, so I cannot respond to this question.

• *Is the discussion on exposure assessment methods (Section 2.3) accurate and complete? If not, what additional information needs to be included?*

I have no relevant expertise, so I cannot respond to this question.

• *Is the discussion on personal exposure (Section 2.4) accurate and complete? If not, what additional information needs to be included?*

I have no relevant expertise, so I cannot respond to this question.

• *Is the discussion on copollutant correlations and potential for confounding (Section 2.5) accurate and complete? If not, what additional information needs to be included?*

As an atmospheric chemist, I am not surprised by the low correlations reported in Figures 2.1 and 2.2, since each pollutant is measured over a time interval different from the ozone measurement; the results in these figures do not provide much useful information. At a minimum, these correlations should be conducted between measurements of pollutants over the same time period; were this done, I would expect substantially larger correlations in some cases.

Overall, this section does not adequately support the stated conclusion that "… confounding of the relationship between ambient ozone exposure and a health effect by exposure to CO, SO_2, NO_2, PM_{10}, or $PM_{2.5}$ is less of a concern for studies of the health effects of ambient ozone exposure compared with studies of the health effects related to exposure of other criteria air pollutants."

• Is the discussion on interpreting exposure measurement error for use in epidemiology studies (Section 2.6) accurate and complete? If not, what additional information needs to be included?

I have no relevant expertise, so I cannot respond to this question.

Appendix 9 – The Role of Tropospheric Ozone in Climate Effects

• Is the discussion on ozone impacts on radiative forcing (Section 9.2) accurate and complete? If not, what additional information needs to be included?

The discussion in Section is complete and reasonably accurate. The evidence for a causal relationship between tropospheric ozone and radiative forcing is very robust. This section relies on IPCC's AR5 report, and quotes its best estimate of tropospheric ozone radiative forcing 0.40 ± 0.20 W/m^2 (from 1750 to 2011). However, in my opinion this quantity may be underestimated by more than indicated by the quoted uncertainty. This estimate is based solely on global chemistry-climate model simulations of the evolution of the global ozone distribution from pre-industrial times to the present. As noted in Appendix 9, long-term changes in free tropospheric ozone and upper tropospheric ozone (where RF is particularly sensitive to changes in ozone) are not captured well by models. The models also do not accurately capture the seasonal cycle of ozone, at least in the marine boundary layer (e.g., Derwent et al., 2018), as well as other observed features of the global ozone distribution. The possible underestimate of the tropospheric ozone radiative forcing should be acknowledged in Appendix 9.

• Is the discussion on ozone impacts on temperature, precipitation, and climate related variables (Section 9.3) accurate and complete? If not, what additional information needs to be included?

Section 9.3 reviews the limited research that has been conducted into a variety of extremely complex questions. I know of no additional information that needs to be included. I believe that the overall conclusion - there is "likely to be causal relationship between tropospheric ozone and temperature, precipitation, and related climate variables." – is accurate and supported by the discussion in the section.

<u>Questions from Dr. Tony Cox</u>

Overarching Questions:

I am not an expert in these questions, so I have no response.

Specific questions:

1. *Question: My question is: Can valid determinations of manipulative or interventional causation – that is, how and whether changing exposure would change health risks – be made based on observed associations of the types analyzed in the ISA?*

I have no relevant expertise, so I cannot respond to this question.

2. *Questions: The following questions are intended to help assess the conceptual clarity and meaning of the causal determination categories, and of key conclusions expressed using them, such as those in Table ES-1 (p. E-5) of the Draft ISA.*

I have no relevant expertise, so I cannot respond to this question.

3. *Questions: Based on these spot checks, I have the following questions:*

I have no relevant expertise, and I am not familiar with the science on ozone and health effects so I cannot respond to this question.

4. *Is the biological evidence presented in the ISA to support causal determinations correctly stated, correctly interpreted, relevant for predicting effects of changes in the ozone NAAQS, and up-to-date?*

I have no relevant expertise, and I am not familiar with the science on ozone and health effects so I cannot respond to this question.

5. *Does the biological evidence presented in the ISA provide well-validated scientific information suitable for predicting the effects on public health of changing NAAQS standard for ozone?*

I have no relevant expertise, and I am not familiar with the science on ozone and health effects so I cannot respond to this question.

6. *Is each of the causal determinations summarized in Table ES-1 (especially those labeled "causal relationship" or "likely to be causal relationship") the only possible causal determination conclusion that is justified by, or consistent, with current scientific evidence? Could different causal determinations be equally well justified (or better justified) by the information presented, or by the totality of current scientific evidence?*

I have no relevant expertise, and I am not familiar with the science on ozone and health effects so I cannot respond to this question.

7. *Are there changes in the design, analysis, selection, or interpretation of individual studies or in the ISA's processes for interpreting and summarizing them that would improve the validity, credibility, and transparency of the ISA's scientific reasoning and conclusions?*

I have no relevant expertise, and I am not familiar with the science on ozone and health effects so I cannot respond to this question.

Questions from Dr. Mark Frampton

1. Change in causality determination for short-term cardiovascular effects since the 2013 ISA.

I have no relevant expertise, and I am not familiar with the science on ozone and health effects so I cannot respond to this question.

2. Metabolic effects, new determination of "likely" for both short- and long-term exposure.

I have no relevant expertise, and I am not familiar with the science on ozone and health effects so I cannot respond to this question.

3. Change in causality determination for total mortality since the 2013 ISA.

I have no relevant expertise, and I am not familiar with the science on ozone and health effects so I cannot respond to these questions

Questions from Dr. Sabine Lange

1. *It has been established that associations found in an epidemiology study can be due to: causation, bias, chance, and/or confounding. If the concept of statistical significance is not useful in epidemiology studies, then how do the study authors/EPA rule out that chance has caused the observed association?*

I can respond to this question as a scientist with extensive experience in interpretation of results based on statistical significance, but not as an epidemiological expert. The concept of statistical significance is useful in interpreting the results of most scientific studies, but it is not of yes-or-no utility. Recent literature (e.g., Amrhein et al., 2019; Hurlbert et al., 2019) emphasizes the importance of not using statistical significance in a dichotomous manner, e.g. to decide whether the results of an analysis rules in or out any particular cause of an observed association. It is recommended to simply present *p*-values without label or category. With regard to the present question, the results of any particular epidemiology study can provide an estimate of the probability that chance has caused the observed association, but that probability can never be reduced to zero. The same statement can be generally applied to the results of all studies that attempt to understand the cause of an observed correlation or association.

2. *Am I correct in understanding that the intention of ozone case-crossover studies is to compare the ozone concentrations on a day when a health effect occurred for a person, to the ozone concentrations on a day when that health effect did not occur for that person?*

I have no relevant expertise, so I cannot respond to this question.

3. *If so, then it would be important that some other factor (not related to ozone) did not prevent the health event from occurring on a control day. These studies often use days before and after the health event as control days, but for mortality studies (such as Di et al., 2017), how can a day after death be used as a control day? It doesn't matter what the ozone concentrations are after a person's*

death, that person would not be able to respond to that concentration. How should we interpret case-crossover studies that use control days after the event (particularly mortality) occurred?

I have no relevant expertise, so I cannot respond to this question.

4. *What is the importance of dose-concordance in establishing the biological likelihood of ozone-mediated effects occurring at relevant exposure concentrations in humans? Particularly in the context of known dose information about ozone: total inhaled dose includes concentration, exposure time, and exercise duration; Hatch et al., (2013) have shown that humans and rats that are exposed to ozone at rest achieve similar alveolar ozone doses, and that humans exercising at 5-times a resting ventilation rate achieved an ~ 5-times higher alveolar ozone dose; and that ozone concentrations are 2-10 times lower indoors where people spend most of their time.*

I have no relevant expertise, so I cannot respond to this question.

5. *Is there evidence that the animal models used to assess ozone effects (largely rats, mice, and non-human primates) are more, less, or similarly sensitive to ozone-mediated adverse effects compared to humans, at approximately equal inhaled doses?*

I have no relevant expertise, so I cannot respond to this question.

6. *In the absence of a causality diagram to direct the choice of variables to control in an epidemiological study, how can we judge whether a study has appropriately controlled for confounders, and has not inappropriately controlled for colliders (which can open up pathways between variables that otherwise would not be connected) or mediators (and thereby controlled away the effect)?*

I have no relevant expertise, so I cannot respond to this question.

<u>Questions from Dr. Corey Masuca</u>

Appendix 1 Atmospheric Source, Chemistry, Meteorology, Trends, and Background Ozone

1.3.1 Precursor Sources

Are there not other chemicals besides CO and CH4 that also are contained in the precursor mix of ozone formation with its rapidly forming and degradation in the atmosphere?

This introductory paragraph also lists volatile organic compounds (VOCs), along with CO and methane. From the perspective of urban ozone pollution, the VOCs are more important than CO and methane as the "fuel" for ozone formation.

Does the singling out of these two constituents of the ozone "cocktail" significant as push toward climate change/global warming instead of just evaluation ozone formation?

My interpretation of this paragraph is that it serves as the foundation for discussion of both ozone formation in urban and rural areas and climate change/global warming.

1.3.1.2.1 Global Methane

Again, is a teasing out/focusing on CH4 important in discussing the virtual "cocktail" of chemicals that may be associated with ozone formation/degradation?

Methane is one of the few species in the "cocktail" of chemicals (carbon monoxide is another) that has a lifetime long enough for it to be transported on intercontinental distances, and thus can directly affect the photochemical production of ozone within the US. However, methane plays only a minor role of in urban photochemistry.

From a US ozone pollution perspective, transport of background ozone into the US makes a major contribution to observed ozone concentrations in urban and rural areas, even during episodes when the NAAQS is exceeded. As discussed in Section 1.3.1.2.1, model studies indicate that these background ozone concentrations increase when methane increases.

However, the dependence of global tropospheric ozone concentrations on methane concentrations has been quantified only by chemistry-climate model simulations. That dependence is expected to be critically dependent upon the model-derived global NOx concentration distribution, and these model simulations are quite sensitive to parameterizations of many physical processes within the models. The parameterizations have been tested by observation-model comparisons only to a limited extent, so their success in realistically simulating the physical processes remains uncertain. Finally, the NOx concentration distribution is poorly characterized from the limited measurements available, and the measured concentrations are often at or below the detection limit of the instruments making the measurements. Testing the simulations of the global NOx concentration distribution is critically needed for improving chemistry-climate model simulations of the global distribution of tropospheric ozone concentrations, and the role that methane plays in determining that distribution.

In summary, the role that global methane plays in the global tropospheric ozone budget is derived from chemistry-climate model simulations. Simulations by different models generally agree, but the models generally use similar parameterizations of critical physical processes. Thus, in my opinion, increasing methane may increase global ozone concentrations, but due to model uncertainties that relationship is uncertain. Conceivably increasing methane may decrease, rather than increase, global ozone concentrations.

1.3.1.2.2 International Emissions of Ozone Precursors

This section focuses on international transport of ozone precursors.

What about local/state/regional transport of ozone precursors?

From the perspective of US ozone NAAQS exceedances, local/state/regional transport of ozone precursors is more important than international transport of ozone precursors. Regional photochemical modeling designed to simulate local and regional ozone pollution events simulate all local/state/regional

processes, including precursor emissions and transport. To the extent that these simulations are physically realistic, the local/state/regional transport of ozone precursors are properly considered.

As mentioned above, transport of background ozone into the US makes a major contribution to urban and rural ozone concentrations. Accurate simulation of international emissions of ozone precursors and their transport is critical for successful chemistry-climate model simulations of global ozone concentrations.

1.3.1.3.2 Biogenic Volatile Organic Compounds (VOCs)

It has been stated that biogenic VOCs and contributions are greater than anthropogenic sources (i.e., motor vehicles).

Is there greater confidence in using models and remote sensing (both with relative degrees of uncertainty) to estimate biogenic ozone source contributions that vehicle emissions estimates (manufacturing vehicle emission standards and testing), in making this assessment?

Estimating VOC emissions, whether biogenic or anthropogenic, is a highly uncertain undertaking due to:

- Multitude of sources (anthropogenic: on-road and off-road vehicles, oil and gas production, solvent use, wild and prescribed fires; biogenic: very many species of vegetation.)
- Multitude of species emitted (literally 100s of different anthropogenic species, and many known and unknown biogenic species).
- High temporal variability on wide scales (diurnal patterns of activity and temperature, seasonal patterns of both biogenic and anthropogenic emissions, longer-term changes, such as evolution of vehicle fleet, industrial processes, and land-use patterns).
- Emissions dependence on details of source use (emissions depend on make and model of vehicle, operation of vehicle, species of plant, etc. etc.)
- Many other sources of uncertainty.

A great deal of effort has been expended to incorporate these details into emission inventory estimates, but large uncertainties remain. Further, any inventory is quickly out of date due to changes in source magnitude and character. Thus, it is difficult to judge whether the biogenic or the anthropogenic emissions are estimated with greater confidence. The remote sensing of formaldehyde by satellite does provide added information regarding biogenic emissions. A similar tool is not available for anthropogenic emissions, but considerably more effort has been invested in quantifying anthropogenic VOC emissions through source characterization studies. However, the satellite estimates have their own uncertainties, and oxidation of anthropogenic VOCs also produce formaldehyde, so the interpretation of the satellite data is often ambiguous in large urban areas. On a national or large regional scale, uncertainties in total annual anthropogenic and biogenic emissions may be as small as a factor of two or even less, but on smaller spatial or temporal scales, or more detailed speciation, the uncertainty is likely significantly larger, and will vary widely depending on the particular emissions in question.

1.4 Ozone Photochemistry

With the advent of monitoring for speciated compounds including PAMS and Near-Road Monitoring (NOy), should there be further discussions about the individual chemicals gleaned from the specialized monitoring.

I see no need for further discussion of any additional individual chemicals. PAMS provides a wealth of information regarding specific VOC species in major US urban areas, but this is detailed information that would not add substantially to Section 1.4, as the aim of this section is to give an overview of recent developments in our understanding of the atmospheric photochemistry that produces ground-level ozone. I do not believe that the PAMS data would be relevant to either the high wintertime ozone observed in western oil and gas production regions, or the role of halogen chemistry in the ozone budget. The Near-Road Monitoring program is designed to collect data at the locations that are most highly impacted by vehicle emissions. These data are needed for transportation system planning, environmental impact assessments, and exposure assessments in health studies. The species monitored are the same as in other urban monitoring programs, so no additional information regarding individual chemicals is obtained.

1.5 Inter-Annual Variability and Longer Term Trends in Meteorological Effects on Anthropogenic and US Background (USB) Ozone

While temperature, wind patterns, cloud cover, and precipitation are highlighted as very important variables in ozone formation, does topography play a role (such as in Birmingham where summertime pollutants are trapped in a "mountainous bowl?"

Topography does play a major role in ozone formation. When air is trapped over an urban area due to surrounding mountains, ozone produced over hours accumulates in that trapped air. The most notable example is the South Coast Air Basin in California, which contains the Los Angeles urban area. It is surrounded on three sides by mountain ranges, with the Pacific Ocean to the west. The interaction of the land-sea breeze circulation with the mountains leads to air circulating over the basin while accumulating ozone, which contributes to Los Angeles recording the highest summertime ozone concentrations in the US. It is very difficult to accurately model the wind fields within the Los Angeles basin, due to the relatively light meteorological forcing and the interactions of the wind with the topographical features. This difficulty contributes to the uncertainty in the photochemical modeling for that region. I am not familiar with Birmingham's topography or air quality, so I cannot comment specifically on that area.

Are there any independent effects on formation formation due to relative humidity?

Relative humidity determines the amount of water taken up by ambient particulate matter, which in turn affects their interaction with the gas phase species involved in ozone photochemistry. Absolute humidity is also important, as water is a reactant in some of the important photochemical reactions that are involved in the formation and destruction of ozone in the atmosphere. Photochemical models incorporate these effects into their photochemical mechanisms.

Appendix 2 Exposure to Ambient Ozone

2.3 Exposure Assessment Methods

While monitoring, including fixed, ambient monitors and personal and microenvironmental monitors are highlighted, what about remote sensing? Biological sampling in blood or tissue?

Remote sensing of ozone is possible with expensive research grade equipment. Even if the difficulties (i.e., laser radiation that is not eye safe) of deploying those instruments in a populated area, their range is limited to a few km. I cannot foresee any role for remote sensing of ozone in exposure assessment studies.

I have no relevant expertise regarding biological sampling in blood or tissue.

2.3.2.1 Spatial Interpolation

While attempting to quantity concentrations at locations and areas between concentration points is included under 2.3.2 Modeling, many of these exact same methods (i.e., data averaging, IDW, and kriging) are also utilized for Monitoring data shortcomings.

2.4.1 Time-Activity Data

Is it possible that ozone exposure through time-activity data may be reduced due to temperature alone, as more people tend to avoid time spent outdoors in the summers during extremely warm/hot/humid, stagnant days which are oftentimes conditions for greater ozone formation?

I have no relevant expertise regarding outdoor activity as a function of outdoor temperature.

Miscellaneous Question(s)

Due to exposure to ozone being disproportionate for disparate (i.e., lower income, children), should this be emphasis in a this section, in lieu of regression analysis confounding/covariate in epidemiological studies for low(er) SES?

I have no relevant expertise, so I cannot respond to this question.

<u>Questions from Dr. Steven Packham</u>

Question 1: When a causal relationship is conclusive to a high degree of scientific certainty as it is in this case, should this take precedence over causal inference when drafting a NAAQS ISA?

I have no relevant expertise, so I cannot respond to this question as an expert; however, to a non-expert the answer is obviously, Yes.

Question 2: Given evidence available from controlled human exposures substantiating causal relationships with a number of physiological responses, including beneficially confounding interactions of ozone on PM clearance, should Sub-section ES4.1 Health Effects in the Draft's Executive Summary, and the entire Integrated Synthesis section of the Draft be rewritten?

I have no relevant health effects expertise, so I cannot respond to this question. However, as an atmospheric chemist, I would dispute details of the second Background Statement of Fact. A human taking about 20 breaths of about 1 liter volume per minute for 6.6 hours with 75 ppb ozone would inhale about 1.3 mg of ozone, not 1,362 mg. However, that 1.3 mg of ozone corresponds to about 160 million trillion ozone molecules, not a mere 2,439 trillion.

Question 3: Looking ahead, do you think toxicology, clinical human studies, and biomedical research disciplines should be given more explicit and balanced consideration in the development of the present, and future, O3 ISAs with the objective to validate causal relationships and determine hourly inhalation dosage rates for adverse inflammatory responses in pulmonary tissues?

Again, my limited familiarity with pulmonary physiology and inhalation toxicology means that I cannot respond to this question.

References

Almaraz, M., et al. (2018). Agriculture is a major source of NO_X pollution in California, *Science Advances*, 4(1) eaao3477, DOI: 10.1126/sciadv.aao3477.

Amrhein, V., S. Greenland and B. McShane (2019). Retire statistical significance, *Nature*, *567*, 305-307, https://www.nature.com/magazine-assets/d41586-019-00857-9/d41586-019-00857-9.pdf.

Derwent, R. G., D. D. Parrish, I. E. Galbally, D. S. Stevenson, R. M. Doherty, V. Naik and P. J. Young (2018). Uncertainties in models of tropospheric ozone based on Monte Carlo analysis: Tropospheric ozone burdens, atmospheric lifetimes and surface distributions, *Atmos. Environ., 180,* 93–102, doi.org/10.1016/j.atmosenv.2018.02.047.

Hurlbert, S. H., Levine, R. A., and Utts, J. (2019). Coup de grâce for a tough old bull: "Statistically significant" expires. *The American Statistician*, *73*(sup1), 352-357.

Jaffe D. A., et al. (2018) Scientific assessment of background ozone over the U.S.: Implications for air quality management. *Elem. Sci. Anth.,* 6 56 doi.org/10.1525/elementa.309.

Langford, A. O., et al. (2017), Entrainment of stratospheric air and Asian pollution by the convective boundary layer in the southwestern U.S., *J. Geophys. Res. Atmos., 122*, 1312–1337, doi:10.1002/2016JD025987.

Miller, C. A., G. Hidy, J. Hales, C.E. Kolb, A.S. Werner, B. Haneke, D. Parrish, H.C. Frey, L. Rojas-Bracho, M. Deslauriers, B. Pennell, J.D. Mobley (2006) Air Emission Inventories in North America: A Critical Assessment, *J. Air & Waste Manage. Assoc., 56*, 1115–1129.

Oikawa, P. Y., et al. (2015). Unusually high soil nitrogen oxide emissions influence air quality in a high-temperature agricultural region, *Nature Communications*, 6:8753 | DOI: 10.1038/ncomms9753.

Parrish, D. D., Young, L. M., Newman, M. H., Aikin, K. C., and Ryerson, T. B. (2017) Ozone Design Values in Southern California's Air Basins: Temporal Evolution and U.S. Background Contribution, *J. Geophys. Res.-Atmos.*, *122*, 11166–11182, https://doi.org/10.1002/2016JD026329.

Parrish, D. D. and C. A. Ennis (2019). Estimating background contributions and US anthropogenic enhancements to maximum ozone concentrations in the northern US, *Atmos. Chem. Phys.*, *19*, 12587–12605, https://doi.org/10.5194/acp-19-12587-2019.

Dr. Lorenz Rhomberg, Gradient

Questions from Dr. Cox

Dr. Cox's set of questions parallels to a large extent the set he posed in regard to the $PM_{2.5}$ ISA. To a large extent, the questions are posed as methodological quandaries rather than as questions about specific instances of application, and those questions are therefore in need of similar responses in the $PM_{2.5}$ and the Ozone case. Accordingly, my responses to his questions on the $PM_{2.5}$ ISA should also be taken to apply here.

As was the case with the $PM_{2.5}$ questions, Dr. Cox raises an overarching fundamental concern about the sufficiency of how causality is understood and its properties characterized and applied, accompanied by a long set of very particular questions about whether these overarching concerns affect many particular details of the conclusions the document tries to justify. To address each of the particulars fully would require a quite massive dissertation on the general questions of how causality can be conceived of, recognized, and applied to support a regulatory analysis, and then a large further set of dissertation chapters – each one quite a task in itself – to apply this general discussion to its impact on all the particular specific inferences Dr. Cox asks about.

I fear that the time and effort available to me as a quick-response consultant does not allow such a deep dive. (It would take a lot of study for anyone not already focused on these issues as a main career path.) I am forced to try to address the thrust of the big questions to the extent I am able with my existing understanding and then briefly comment on selected application questions.

As I suggested in the comments on the $PM_{2.5}$ ISA, the issues raised are legitimate concerns, and what is needed is not quick responses from an array of commenters to a lot of particular instances of the larger questions, but rather a full taking on of a deeper look into causal inference as it can and should be applied in the evaluation of criteria air pollutants. This seems a topic for an appropriately chosen and charged NAS committee, with the time and expertise to bring epistemological principles, insights, and approaches into a general approach to assessment of the roles of particular pollutants in affecting public health, based on the kinds of data that are actually available.

Collectively, the questions raise the issue as to whether the interpretation of studies as evidence of causal influences of the agents under review with regard to potential health impacts have been rigorously pursued. Dr. Cox's questions raise the problem of whether appropriate distinctions are being drawn in the ISA analyses between observations of an agent's apparent association with toxicity outcomes of interest (which association one might *hypothesize* to be attributable, at least in part, to an actual causal role of the agent vis-à-vis the outcome) and a conclusion that the agent does indeed have such a causal role. There is the further issue of distinguishing the general ability to influence a toxicity outcome in some way under some circumstances (a qualitative property of the agent vis-à-vis its potential targets, potentially contingent on other factors, but asserted as the existence of potential for action divorced from particulars), and the inference that the agent has indeed acted causally in particular observed circumstances. The key question for the ISA application is really the reverse of this last one: given a set of particular observations about outcomes of interest and associated exposures to the agent, what kinds of associations – and what analytical approach to those associations – can justify an assertion

that the agent operated as a causal influence in contributing to the bringing about of those outcomes? This needs to be distinguished from the further question as to whether the agent would be expected to exert causal effects in *other* circumstances that will not share all the particulars of the observed case – that is, is the assertion of a causal effect seen in some studies can be generalized as a basis to predict that causal effects would be expected from other exposures? This in turn is distinct from the question of whether, if such qualitative prediction is deemed possible, the *magnitude* of the effect can be measured from observations to predict the magnitude of effect in hypothetical future exposure settings where other circumstances might differ.

My overarching reaction to these large questions is that Dr. Cox's concerns have merit. The treatment of "causation" in the ISA tends to be fuzzy – fuzzier in some places more than in others – about the distinctions noted above. If patterns of association are plausibly explained by underlying causation, this is taken as sufficient evidence for such causation (when one should actually be comparing the hypothesized causative actions against other competing explanations for the patterns), and if such causation is inferred, it is taken to be universal, applying to other settings on the usually poorly stated (much less justified) presumption that the causation is universal and largely independent of other circumstances. The measures of association are too easily taken as measures of the magnitude of cause, and these magnitudes applied to other settings without due consideration of how they may be contingent on differing circumstances.

This said, it should be clear that there is no ready "fix" to the issues raised. There certainly are procedures that can help identify and disentangle the various influences and contingencies -- and their use should be encouraged – but none will be absolutely definitive. There is no ideal analysis that can unambiguously characterize casual properties and their functional dependence on circumstances, such that an unambiguous and unquestioned prediction of the consequences of future exposures can be estimated without risk of tripping on one or several of the challenges that Dr. Cox's questions raise.

Clearly, one should attempt to identify and adjust for confounders. Directed graph analyses can help sort out questions as to which variables act directly *versus* through their effect on other variables. Experiences with outcomes that differ in their input patterns, along with controlled experiments, can help identify what factors can or cannot reasonably be expected to influence or interact with one another regarding outcomes. One should think through inferences to see where simplifying assumptions might be covering important influences.

The approach, then, should be to be aware of the pitfalls and challenges, and to beware of being naïve about the difficulty of sorting out the existence of some causal effects or about the challenges of asserting general properties from particular observations and asserting that they would apply, independent of other influences, in future settings that one wishes to predict to address the controls on exposure that would "protect public health with an adequate margin of safety." Just because a potential interpretation problem cannot be definitively eliminated, it does not mean that the information is meaningless or that there is no useful evidence about potential effects. But at the same time, the inability to eliminate a potential problem that cannot definitively be demonstrated does not mean that the problem does not exist or can simply be ignored.

It seems that much of the problem arises because of the tendency to think of a "cause" as both necessary and sufficient for the outcome. "Necessary" in that any effects seen are presumed to have come from the

potential cause of interest (and not from other influences), and "sufficient" in that that cause alone is all that is needed, with no contingency on other factors as contributors or facilitators. This simple notion of "cause" will not really do for the kinds of effects pollutants can have on health.

Most pollutant effects on health will be INUS (Mackie 1974) – individually insufficient but (nonredundantly) necessary parts of a process that collectively is unnecessary but sufficient.

- That is, "insufficient" because other elements need to be present (behavior causing uptake, individual susceptibility is sufficient, etc.); but "necessary" (process will not happen without it, i.e., the "efficient cause" or "that without which"); "sufficient" because together the necessary element and its other components are enough to result in the effect; but "unnecessary" in that this collection of components is not the only way to get the effect (there are other causal pathways that can get to the same end result).
- Thus, the empirical association of a putative cause with effects is not dispositive for concluding it has a "causative" role, since (1) other, separate causative processes could have been responsible for the effects of interest (such that seeing the effect does not unambiguously identify the cause of it), and (2) the role of the putative cause – even if true – is contingent on other enabling features of the set of conditions (such that failing to see the effect after the putative cause could be explained by other needed conditions being missing).

What we usually intend to mean by something being "causative" is that it is an "efficient cause" in Aristotle's sense – the actor that imposes on other elements and, as a consequence of that instance of imposition, precipitates a change in their states. It is important to distinguish between (1) the assertion that this is an ability that an agent may exert (in some circumstances, i.e., a property of the agent's potential with respect to the states it might alter) and (2) the assertion that the agent has in fact led to a specific observed alteration of state (an assertion that, but for the presence of the agent, the particular alteration would not have occurred). These are both distinct from (3) the assertion that, whenever the agent encounters an appropriate target, it will act to effect its change – that is, the cause is by itself "sufficient" (not dependent or contingent on other circumstances; and (4) the assertion that an observed change can only be due to the specific cause (that is, that the cause is "necessary" for the effect, such that the observation of the effect ensures that the cause operated, with no other means available to bring about the effect).

Broadly speaking, there are two kinds of complications that arise: (1) the existence of potential confounders and varying influencers of outcomes, and (2) the further potential complexity arising from patterns of how various factors interact. The most obvious is the role of potential confounders. Several of the criteria air pollutants cause effects on the same endpoints, and the levels of these pollutants is highly correlated in space and time. Any analysis that does not adjust for such exposures will falsely attribute the entire effect to the one pollutant chosen as the "exposure" variable. In addition, non-pollutant factors can have a role, notably, the potential impact of differing weather (especially temperature) which can plausibly play an independent role, modify the impact of pollutant levels, and be correlated with those levels. To an extent, confounding can be mitigated by adjusting for the levels of the potential confounders in the analysis, but this can only go so far in eliminating the problem. One must recognize the factors in the first place, and many influential factors that ought to be adjusted for may be unrecognized or have insufficient data.

More than this, any adjustment itself makes untested assumptions about independence of the several influences and how they might interact to produce effects. Even if one has allowed for the correlated appearance of copollutants, to gauge their joint impact requires assumptions about whether they independently affect the outcome variable or somehow interact synergistically. In real populations, the values of the important variables will vary among individuals, and the particular combinations experienced by different people will differ, so what is observed is some kind of average over all the possibilities present in the measured population – averaging which can obscure the role of potentially important interactions. Any tractable analysis must make assumptions – usually untestable – about how observed averages over person-to-person differences in levels, differences in combinations of exposures, differences in susceptibility, differences in the timecourse of exposure, and so on combine to yield an effect. If all the effects are independently acting and linear, then one can sum up the effects and express each as the average effect from average exposure (since for linear functions, the averaged output for variable inputs is equal to the function's value for the average input), but if they are not, it will depend on how they are not.

Clearly, these are not new concerns – they have been central to discussion of epidemiological methods through the history of that discipline. As I have argued, the key is to avoid being paralyzed by them but not to do this by ignoring them. One has to attempt to gauge the extent to which the conclusions from simplified but tractable analyses might be skewed by known or reasonably suspected complications, and then treat these insights as insights into uncertainty, to be factored in when one applies the analyses to decisions. Tools exist to help sort out the difficulties, and they should be used, but in the end, forthright expert judgment needs to be employed and (especially in the context of public-policy decisionmaking) the reasoning behind uncertainty characterizations being made explicit.
The way these issues affect the "causality" problem is that the complexities of context and potential interactions are what lead to the difficulty in inferring between the different aspects of causality – ability to affect, responsibility for observations, ability to generalize effect to other settings, ability to measure the impact of an effect, and ability to generalize that magnitude to other settings. They also affect the certainty with which causal determinations can be made, constituting the things that need to be thought through in characterizing uncertainty in any causation assertions.

As to the specific questions, I can offer the following.

Question 1 – It is not possible to make an undisputable, totally sufficient conclusion of interventional causation, because of the INUS nature of the possible causative processes. That is, there is no absolute necessity or sufficiency with which to show such causation by direct means (by giving or withholding the putative cause and seeing or not seeing the effect with absolute consistency). This is not to say that a hypothesis of interventional causation could not be proffered and examined against all the data that would be expected to show such an effect, paired with alternative explanations for how the effect pattern might arise from other plausible causes, and their comparative support compared. But it would have to acknowledge the possibility of other causes for the observed effects and of contingency of the interventional effect on other circumstances. This question of interventional causation, even if solved, does not release one from the challenges of sorting out what sets of conditions may contribute as well.

Question 2a – "a formal causal framework" implies that a path to certain determination is available, and I do not think one is. That said, one could establish a more explicit framework for assessing causation,

and the NAS committee suggestion above would be a good way to pursue such general guidance – probably better than trying to write one in to an already existing ISA.

2b – No, as noted above, the naïve notion of causation sort of implies both necessity and sufficiency, but a more appropriate INUS view of how causes enter into health effect influence shows that the question is more complex.

2c – No, existence of a causal possibility does not guarantee that it would operate in any particular way. This is a further inference about its generalizability. The assertion that a cause operates independent of other varying potential influences is itself a strong conclusion that would be hard to defend in any absolute terms. So it is a question of working through how an effect seen in one setting may or may not show up in others, with an assessment of any hypothesized independence-from-setting properties needed.

2d – Yes, it can be incorrect. It cannot be surely correct, so the question is how usefully to characterize the uncertainty and also how to act in the face of such uncertainty, given that failure to prove the correctness cannot be avoided, but it also does not prove that the causal process does not operate.

2e – I have argued that this could be improved, though it is hard to spell out a remedy in the context of these questions.

2f – there are methods to try to uncover contingencies on other factors, by measuring them and using directed graph analysis. But even these depend on the possible factors being noted and measured, and on other factor being absent, and on assumptions about independence of factors or lack of modification of risks by variable factors from case to case being true, and there will always be possibilities that cannot be addressed by available data and possible analyses. The world is always more complex than any feasible model, and the differences can affect conclusions, so one is always assuming that the needed simplifications do not miss something that is key, with plausibility arguments coming in where data are insufficient to give insight.

2g – I have argued that a more rigorous conception of causation is indeed needed. Bradford Hill considerations and other such approaches are at best guides for fuller thinking through, they are not tests to be passed.

2h – as argued above, no.

2i – as argued above, no.

2j – as argued above, no.

2k – unclear, partly because the meaning of each category could be tweaked. What is clear is that the categories do not seem well suited for making the distinctions between existence of some causal property, the degree of independence of that property from other circumstances, and the possibilities and limits of generalization.

2l – see 2k

2m – this is a matter of being clearer about hones standards.

Question 3 – In the time available, I am not able to chase down and document a basis for all these particulars, and they are more about reviewing the results than commenting on the technical method. My answers above will have to be used to address these questions.

Mackie, John Leslie (1974). The Cement of the Universe: A Study of Causation. *Clarendon Press.*

Dr. Sonja Sax, Ramboll

Questions from Dr. Tony Cox

Overarching Questions (*in italics*)

1. *Is the scientific information provided by the ISA clear?*

    a. *Is it clear how the ISA's causal determination conclusions can be tested, and either verified or refuted (or left undecided), by observations?*

    In my opinion, the ISA could be clearer in many aspects, including how studies are selected for inclusion (or exclusion), and how the evidence is weighed and specifically used to reach causal conclusions. I generally agree with many of Dr. Cox's suggestions for improving the process to make it more objective. Given the same information, I believe, that different people (or different groups) could come to very different conclusions and causal determinations based on current classification descriptions.

    In the ozone ISA, this is exemplified by the "down grading" of the classification for two specific endpoints – cardiovascular effects and mortality from ozone exposures. Given similar evidence in the prior ozone ISA, EPA concluded that both of these endpoints were "likely causal," whereas now EPA has concluded that the evidence is "suggestive." Several papers published by me and my colleagues, which questioned the original EPA classification for cardiovascular effects are more in-line with the current conclusions by EPA. On the other hand, EPA has concluded that evidence supports a "likely causal" classification for metabolic effects, for which EPA did not provide any classification previously. There are many issues with this new category of health outcomes, and with the way in which EPA evaluated the evidence and made its causal conclusion determinations. For example, being a relatively new health outcome category, which greatly overlap with other health outcomes (i.e., cardiovascular effects) it is unclear how the evidence should be evaluated, including whether studies have sufficiently accounted for the numerous confounders that could contribute to the findings. Furthermore, it is less clear how the animal findings (which appear to be the evidence that EPA finds most compelling) should be interpreted. That is, whether the various downstream effects observed at high levels of exposure would be applicable to development of metabolic disease in humans.

    In a 2013, my colleagues and I published a paper that presented recommendation for improvements to the NAAQS causality framework, including many areas that could benefit from greater clarity. In particular, we use the prior ozone ISA as a case study to highlight areas where EPA could improve the overall evaluation of health effects. This paper is relevant to the current ISA as many of the limitations associated with the prior ISA are still present.

[See Goodman JE, Prueitt RL, Sax SN, Bailey LA, Rhomberg LR. **Evaluation of the causal framework used for setting national ambient air quality standards.** Critical Reviews in Toxicology 2013;43(10):829-49.]

For example, the scientific evidence provided in the ISA generally represents a summary of the findings from selected literature that EPA deems to be most relevant. In general, I understand that the literature that EPA is tasked to review is quite vast, however, the ISA could benefit from being clearer on how the literature is selected (i.e., more specifics on search criteria used, screening on literature, literature included and excluded, and reasons for inclusion and exclusion). This would provide greater transparency related to this process.

One critique in our paper that EPA has partially addressed in the current ISA, is that there previously was a lack of guidance regarding how study quality in considered. In the current ISA, a new addition is some guidance on what constitutes a high-quality study in each of the study domains (epidemiology, toxicology, etc.), which is essentially repeated as an Annex to each Appendix for each health outcome category. However, it is still unclear how the guidance is applied to each individual study, as the study summaries presented little, if any, information related to the study quality or even the strengths and limitations of the studies. EPA could be clearer on whether the study quality guidelines were used to include/exclude studies that fit/or did not fit these study quality criteria. That is, are all the studies that EPA presents considered to be "high" quality studies? I have not verified whether the studies summarized comply with the study quality criteria that EPA presents. While this shows progress in the right direction, EPA needs to take it one step further and present information on how the guidelines were used in the evaluation of individual studies.

In addition, the ISA causality framework would also benefit from some clarity regarding what each causal classification represents. Specifically, is the classification meant to identify the strength of the evidence for a given level of exposure (e.g., below the current NAAQS) or is it simply to identify whether there is evidence enough of expected harm at any level of exposure? That is, does the ISA causality framework represent simply a hazard identification rather than identification of harm at a certain level of exposure. EPA has made some progress on this point by identifying relevant literature to be evaluated using PECOS, which "defines the parameters and provides a framework to help identify the relevant literature to inform the draft 2019 Ozone ISA statements." However, in the PECOS statements EPA specifies an exposure range to be considered. i.e., "0.4 ppm or below for humans, 2 ppm or below for other mammals" only for experimental studies, and for epidemiological studies only specifies "ambient concentrations of ozone," presumably including concentrations above and below the NAAQS. In general, EPA could be more explicit in detailing the distinction between hazard assessment and risk assessment (for example, in the actual definitions for the causal classifications).

b. *Do the concepts and terms used to express key scientific conclusions in the ISA, especially the causal determination categories, have clear scientific meanings (e.g., unambiguous operational definitions)?*

The operational definitions of the causal classifications leave a lot of room for subjective interpretation. For example, for a causal conclusion EPA notes that "the pollutant has been shown to result in health effects in studies in which chance, confounding, and other biases could be ruled out with reasonable confidence." In this case it is unclear what EPA means by "reasonable confidence." What is reasonable to one person may not be sufficient for another.

c. *Is it clear and generally understood and agreed what the key conclusions mean? Specifically, are the causal determination categories used to communicate key conclusions unambiguous and well defined?*

See answer above

d. *Do those who read the ISA have a shared, unambiguous understanding of what its key scientific conclusions (i.e., causality determinations) imply about how or whether changes in ozone air pollution would change public health outcomes?*

This is where further clarity is needed. Are the causal classification conclusions suppose represent a hazard identification only? Or should they also inform whether harm occurs at a certain "relevant" level of exposure? This is unclear in the ISA, where for example, PECOS statements define a level of exposure for experimental studies, but not for epidemiological studies. Because the NAAQS process involved the development of a level of exposure.

2. *Is the scientific information provided by the ISA sound?*

a. *Are its conclusions logically implied by the data and analyses on which they are based?*

Most of the ISA sections that I have reviewed (albeit not in very much detail), including the respiratory section (Appendix 3), the metabolic effects section (Appendix 5), and the cardiovascular section (Appendix 4) as well as the Integrated Synthesis, appear to summarize results from recent studies and provide very general "integration" summaries. There is little with regard to presenting information on study quality or how studies are weighed by EPA when making inferences or drawing conclusions. In general, it is difficult to follow the rationale as presented in the ISA with regards to causal conclusion determinations and there are appear to be inconsistencies in how the evidence is deemed sufficient to select one classification vs. another, such as in the cardiovascular vs. the metabolic outcomes. This highlights the need to have clear protocols and definitions regarding how these conclusions are derived.

b. *Are its conclusions correctly stated and caveated?*

No, it is hard to follow the rationale for the conclusions. For example, although the cardiovascular effects section and metabolic effects section are related, EPA comes to very different conclusions regarding these two endpoints. EPA seems to provide much more discussion regarding the uncertainty and limitations of the cardiovascular evidence, in particular the incongruency between the subclinical results in animal studies and the epidemiological findings. In addition, there appears to be many more studies on the cardiovascular effects, than on metabolic effects and I saw little if any discussion of strengths and limitations of any of the studies in the metabolic effects section compared to the cardiovascular section. I think it is essential that EPA include this information in the text or the tables. For example, if animal studies only included a single exposure dose these studies provide much more limited evidence of an effect because you cannot evaluate a dose-response relationship. For studies that do provide multiple exposures, it is important to consider dose-response, and EPA does not seem to consider this. As noted above, adversity of effect is also an important consideration as transient, reversible effects are less likely to be adverse.

c. *Is it clear than conclusions do not reflect selection bias in the choice of studies relied on? Are its conclusions consistent with other relevant data and studies not included in the ISA?*

I have not conducted a separate literature search or analysis of the data to determine if there is selection bias. I do note that several studies (some of which I am an author of and references provided throughout) are missing from the ISA. More transparency is needed on how EPA selected studies for inclusion/exclusion

d. *Is it clear how studies were selected for inclusion in the ISA, and why individual studies were included or excluded?*

No, I did not see a discussion of how studies were selected or why studies were included or excluded, except for reference to the PECOS statements. I think that more clarity is needed in this regard.

Specific Questions

1. *Question: Can valid determinations of manipulative or interventional causation – that is, how and whether changing exposure would change health risks – be made based on observed associations of the types analyzed in the ISA? I emphasize manipulative and interventional causation (rather than predictive (Granger) causation, but-for causation, epidemiological (attributive) causation, mechanistic causation, etc.) because it is most relevant for policy makers.*

The short answer to this question I think is no. I suspect that Dr. Cox is referring to so called accountability studies, such as his paper [Cox & Popken. 2015. Has reducing fine particulate matter and ozone caused reduced mortality rates in the United States? Annals of Epidemiology, Volume 25 (3), pgs. 162-173], which explicitly address whether reductions in PM and ozone have contributed to reductions in mortality rates. As noted above, EPA does need to make it clear

whether the ISA is simply a hazard evaluation or if it is meant to address the adversity of the effects at a given exposure level.

2. *Questions: The following questions are intended to help assess the conceptual clarity and meaning of the causal determination categories, and of key conclusions expressed using them, such as those in Table ES-1 (p. E-5) of the Draft ISA.*

   a. *A preliminary question: Is this actually a "formal causal framework"?*
   I am unsure of the answer/ or answer is not clear. I generally agree that the current definitions for each classification leave a lot of room for subjective judgement and in general it is not always clear (to me) how EPA weighs the evidence and comes to its final causal conclusions. This could use some revision and clarification.

   b. *Does the ISA's causal determination framework clearly distinguish between necessary and sufficient causation?*

   The short answer to this is no. The ISA does not clearly make a distinction between necessary and sufficient causation as specified by Dr. Cox. This relates to my discussion above regarding clarity on whether the ISA is simply a hazard identification or whether the ISA is meant to identify a level of exposure (as tasked for developing a NAAQS) at which adverse effects are expected to occur. The current ISA and overall NAAQS framework is not clear on this point.

   c. *Does a determination that exposure has a "causal relationship" with a health effect in a population imply that reducing exposure would reduce risk of the health effect in the population, other factors being held fixed?*

   I think again the answer is no. I don't think the data are sufficient to establish this with any certainty. There are too many other factors that influence the development of disease or that contribute to deaths and the processes are so complex that it would be difficult, if not impossible, to assess this with any degree of certainty, particularly at low exposures. The observed associations do not imply causation and it is only by looking critically at all of the data, across all lines of evidence and weighing the strengths and weaknesses of these lines of evidence that we can at least establish plausibility for causation. However, even if causation could be established, given the multiple risk factors for any given disease (respiratory, cardiovascular, etc.) it would be difficult to determine how much reduction in risk from air pollution exposures would influence reductions in disease with any certainty. This is difficult for any risk factor that is evaluated, some which are known modifiable life style choices (smoking, drinking, diet, and exercise) and other which are not (e.g., genetics).

   d. *Can causal determinations be incorrect? (Or, to the contrary, are they performative utterances?)*

   I think that the causal classifications can be interpreted differently based on the current definitions and how EPA presents and interprets the data. That is, they are open to interpretation, particularly for determinations of a "causal relationship." That may be the only

classification that could be considered to be "incorrect" based on different interpretations and weighing of the available data because I don't think that causality can be established with absolute certainty. Other classification determinations are more open to interpretation as to whether they are correct or not.

e. *If causal determinations can be mistaken, then is it clear how uncertainty about which category is correct should be (or has been) resolved in assigning a final causal determination category, as in Table ES-1 p. ES-5) of the ISA?*

No, I don't think it is clear. As noted above, EPA should provide more clarity and perhaps caveat the classifications based on the amount of uncertainty in the underlying evidence.

f. *If causal determinations can be incorrect, then is it clear how observations could be used to test and falsify a given causal determination if it is not correct? For example, is it completely clear how someone can use relevant data to show that a determination of "causal relationship" or "likely to be causal" in the ISA is incorrect, if indeed that is the case?*

No, I don't think that as the framework is laid out by EPA, it could be applied consistently and that the same causal determinations would be necessarily developed by different groups. As noted above, this is exemplified by the change in classification for two key health outcomes (CV and mortality) in the current ISA.

g. *If causal determinations can be incorrect, then is the correctness of each causal determination in table ES-1 formally and transparently evaluated in the ISA? In other words, have formal rules for determining the correctness of the causal determinations in Table ES-1 (p. ES-5) from the data and evidence presented been explicitly stated, applied systematically, and the results documented? (If so, where?)*

No, see my comments regarding the NAAQS framework above, as well as Goodman et al. (2013) reference (above).

h. *Does a determination that an exposure-response (or concentration-response (C-R)) relationship is a "causal relationship" imply that it is entirely causal, with no contribution from incompletely controlled confounding, modeling errors and biases, or other non-causal sources? If not, is there a clearly defined lower bound on how much of the relationship (e.g., how much of the slope of a C-R regression line) must be causal in order for the whole relationship to be classified as causal? (If so, what is it?)*

This is an interesting question, and I think that the way that EPA uses the C-R functions in the risk assessment, assumes that there is a causal relationship between the air pollutant (in this case ozone) and the various health effects, under the conditions specified in the underlying epidemiology model and then EPA often applies this to other populations. However, the epidemiological studies are flawed in that they do not account for uncontrolled or incompletely controlled confounders, and have errors and biases that are not always discussed or caveated. Therefore, it is more likely that there does exist some lower bound of the attributable fraction and it is unclear how much of it is necessary for consideration of a

true causal effect. I think this is somewhat akin to determining a threshold below which effect are unlikely to occur, something that is very likely for ozone based on the lungs protective mechanisms against oxidative damage.

i. Does a determination that a C-R relationship is a "causal relationship" imply 100% certainty that it is causal? If not, is there a clearly specified lower bound on how probable it must be that the relationship is causal in order for it to be classified as causal? (If so, what is it?)

See previous answer. I think as EPA uses it, yes.

j. Does a determination that a C-R relationship is a "causal relationship" imply that it is causal for every member of a population, or might it be deemed "causal" if it is causal for a sensitive subpopulation only? In the latter case, is there a clearly specified lower bound on the fraction of the exposed population for which the relationship must be causal, in order for the whole relationship to be classified as causal? (If so, what is it?)

This is also an interesting question. EPA currently uses C-R functions in its risk assessment based on selected epidemiological studies. The epidemiological studies are based on different population groups (sometimes children, sometimes the elderly, sometimes all ages etc.), therefore I think it depends on the epidemiological study and the population group that the study includes.

k. Are the five categories mutually exclusive?

I don't think they are necessarily mutually exclusive. Again, as defined I think the scientific evidence could be evaluated differently by different people or groups such that they would arrive at different conclusions (or different categories) given the same evidence. See for example our assessment of the cardiovascular effects of ozone and EPA's conclusions in the prior ozone ISA.

Prueitt, RL; Lynch, HN; Zu, K; Sax, SN; Venditti, FJ; Goodman, JE. 2014. "Weight-of-evidence Evaluation of Long-term Ozone Exposure and Cardiovascular Effects." *Crit. Rev. Toxicol.* 44(9):791-822.

Goodman, JE; Prueitt, RL; Sax, SN; Lynch, HN; Zu, K; Lemay, JC; King, JM; Venditti, FJ. 2014. "Weight-of-evidence Evaluation of Short-term Ozone Exposure and Cardiovascular Effects." *Crit. Rev. Toxicol.* 44(9):725-790.

l. Are the five categories collectively exhaustive?

Not necessarily, as there is no counter classification to each of the levels as noted by Dr. Cox, except if EPA selects the last category "not likely to be a causal relationship" to be the counter point to all other classifications – including "causal" and "suggestive" as well as "likely causal."

m. *Can a body of evidence be categorized as "likely to be causal" if the probability of causality based on the evidence is less than 50%?*

Based on EPA's definitions, it is hard to say what percent probability of causality needs to be for the evidence to be considered "likely to be causal."

3.

a. *Questions: Based on these spot checks, I have the following questions:*

i. *Is it clear that the ISA's study selection process has successfully provided a comprehensive, trustworthy, and unbiased selection of the best available science on ozone and health effects?*

As noted above, the ISA would benefit greatly from more transparency regarding the literature search strategy, the selection process for studies included for evaluation, and a summary of why certain studies were included or excluded from consideration.

ii. *Is it clear why results from Moore (2008) are included and cited as "key evidence" but contrary results from Moore (2012) are excluded? More generally, is it clear that study inclusion and exclusion criteria were applied systematically and neutrally to identify and select the best and most up-to-date studies to inform the ISA's conclusions?*

This is an area that does require further evaluation. I agree that EPA should include both positive and negative findings in a balance manner so a not to appear like it is cherry picking.

iii. *Are there other studies that are omitted from the ISA that should be included?*

Although I have not done an independent literature search, it appears that Dr. Cox identified studies that were not included (and EPA did not provide and explanation for exclusion). I note above, and also in comments elsewhere, several publications that my colleagues and I have published that could be considered for inclusion in the ISA.

iv. *Are there studies included in the ISA that should be omitted (e.g., because of uncontrolled confounding, obsolete or incorrect modeling assumptions, conclusions dependent on unverified assumptions, ecological fallacy, lack of causally relevant information, lack of design that can support valid causal inferences, or other methodological problems?)*

Other than general study quality criteria that EPA has identified in an Annex to each of the Appendices, it is still unclear how EPA weighs studies based on study quality. This is an area that EPA still needs to work on and add to it's ISA evaluations. After an assessment of study quality, which would help to identify

methodological issues or biases in certain studies, EPA could develop criteria for excluding studies or at the very least for giving certain studies less weight in the overall conclusions and causal determinations.

v. *Is it clear that the process followed in selecting and summarizing scientific studies in the ISA was sufficient to assure accurate, unbiased, up-to-date, and trustworthy summaries of the relevant scientific literature to inform causal determination judgments?*

vi. *Do you find in the Executive Summary a clear explanation of the extent to which the key evidence supporting the ISA's causal determinations consists of, is sensitive to, or is derived from unverified modeling assumptions, or from modeling assumptions that more recent literature has found to be incorrect or inadequate? Have you found information in the ISA on sensitivity of causal determination conclusions to untested, uncertain, or incorrect assumptions? (If so, where? See Table Annex 6-1, cf p. 6-67 for a discussion of what should be done. Has it be done, and is it clear what the results were?)*

b. *Were the epidemiological studies used to support the causal determinations summarized in Table ES-1 (p. ES-5) and Figure ES-2 (p. ES-6) appropriately designed and analyzed to provide valid scientific information and valid causal conclusions about effects of possible future interventions (rather than just conclusions about historical statistical associations. For these observational studies, were criteria for valid study design and analysis for causal inference (specifically for interventional causation) explicitly stated, systematically applied, and the results transparently presented? (If so, where?)*

c. *Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" adequately controlled for potential confounding and residual confounding by variables such as income and weather variables?*

I did not have the time to conduct an in-depth evaluation of whether EPA applied the criteria for study quality and causal determinations (e.g., Table Annex 6-1) appropriately or correctly. Generally, I found that study summaries for most of the health outcomes I reviewed lacked a clear discussion of the study strengths and weaknesses and there did not appear to be any discussion regarding study quality or how the EPA criteria were applied to individual studies, even for studies that were identified as being "key evidence." This is an area that could be significantly improved. Overall, I think that this limits the transparency of how EPA arrived at the causal conclusions. There are clearly many limitations for all the various studies evaluated and included in the ISA, and these limitations need to be more explicitly discusses and weighed.

I did not have the time or opportunity to review and answer additional questions posed by Dr. Cox, however, I think that some of the answers provided above, partially answer many of these additional queries. Overall, I think that EPA is going in the right direction with developing PECOS statements and providing guidelines for assessment of study quality. It falls short, however, in the implementation and in clearly and transparently showing how these criteria are applied to individual studies.

Questions from Dr. Mark Frampton

1. Change in causality determination for short-term cardiovascular effects since the 2013 ISA.

Question 1: Please comment on the strengths and weaknesses of the epidemiology literature with regard to CV effects of short-term ozone exposure. Are there key studies that are missing? Are the remaining weaknesses, along with the other new evidence, sufficient to justify the change in causality determination?

I believe that the change in causality classification is justified. I refer Dr. Frampton to several studies that were published by me and my colleagues, and the findings from these evaluations are more consistent with the change in the causality determination for CV disease. None of these studies were mentioned or included in the ozone ISA:

Prueitt, RL; Lynch, HN; Zu, K; Sax, SN; Venditti, FJ; Goodman, JE. 2014. "Weight-of-evidence Evaluation of Long-term Ozone Exposure and Cardiovascular Effects." *Crit. Rev. Toxicol.* 44(9):791-822.

Goodman, JE; Prueitt, RL; Sax, SN; Lynch, HN; Zu, K; Lemay, JC; King, JM; Venditti, FJ. 2014. "Weight-of-evidence Evaluation of Short-term Ozone Exposure and Cardiovascular Effects." *Crit. Rev. Toxicol.* 44(9):725-790.

Goodman, JE; Prueitt, RL; Sax, SN; Pizzurro, DM; Lynch, HN; Zu, K; Venditti, FJ. 2015. "Ozone Exposure and Systemic Biomarkers: Evaluation of Evidence for Adverse Cardiovascular Health Impacts." *Crit. Rev. Toxicol.* 45(5):412-452.

Petito Boyce, C; Goodman, JE; Sax, SN; Loftus, CT. 2015. "Providing Perspective for Interpreting Cardiovascular Mortality Risks Associated with Ozone Exposures." *Reg. Tox. Pharmacol.* 72(1):107-116.

2. Metabolic effects, new determination of "likely" for both short- and long-term exposure.

Question 2: Is there sufficient epidemiological evidence of metabolic effects to justify the "likely" determination for both short- and long-term exposures? Are there additional studies that should be considered?

I found the data to be quite limited for this large group of health outcomes, with a lot of overlap between many of these effects and effects that could potentially contribute to cardiovascular disease or even respiratory disease. I think that the large leap that EPA took to conclude that these effects are "likely to be causal" is very premature and the evidence (as presented by EPA) does not appear to justify the classification.

3. Change in causality determination for total mortality since the 2013 ISA.

Question 3: Please comment on the strengths and weaknesses of the epidemiology literature with regard to short-term ozone exposure and total mortality. Are there key studies that are missing? Does the available evidence justify the change in causality determination for total mortality?

EPA seems to be appropriately acknowledging many of the important limitations in the epidemiology literature for ozone and mortality, including potential confounding by other air pollutants, weather and temporal trends, large amounts of unexplained heterogeneity in ozone-mortality effect estimates, and potential for exposure measurement errors or modeling errors. Although total mortality appears to have consistent findings across studies, the cause-specific mortality results are inconsistent, largely null and very imprecise (Figure 6-2).

Questions from Dr. Sabine Lange

Epidemiology Study Questions

1. *It has been established that associations found in an epidemiology study can be due to: causation, bias, chance, and/or confounding. If the concept of statistical significance is not useful in epidemiology studies, then how do the study authors/EPA rule out that chance has caused the observed association?*

   Epidemiological studies cannot be the sole basis for establishing a causal relationship because of the inherent limitations and because, in the case of observational epidemiological studies, you cannot rule out bias, chance and/or confounding with reasonable confidence. The issue is even more difficult when the observed effects are very small and not statistically significant or marginally significant. Because of these limitations it is essential to evaluate all lines of scientific evidence, including experimental evidence (human chamber studies, animal studies, mechanistic studies). By evaluating the consistency and coherence within and across the various scientific lines of evidence, one can obtain a better picture of whether a causal association is more or less likely. More importantly, the evidence may be able to also elucidate levels of exposure at which effects are more likely.

2. *Am I correct in understanding that the intention of ozone case-crossover studies is to compare the ozone concentrations on a day when a health effect occurred for a person, to the ozone concentrations on a day when that health effect did not occur for that person?*

   I believe that is the correct interpretation.

3. *If so, then it would be important that some other factor (not related to ozone) did not prevent the health event from occurring on a control day. These studies often use days before and after the health event as control days, but for mortality studies (such as Di et al., 2017), how can a day after death be used as a control day? It doesn't matter what the ozone concentrations are after a person's death, that person would not be able to respond to that concentration. How should we*

interpret case-crossover studies that use control days after the event (particularly mortality) occurred?

I think that this is a valid question, as you described by selecting days post-health effect this would violate an important epidemiological tenant for assessing a causal relationship – that is, that the exposure must precede the effect.

Experimental Study and Dose Concordance Questions

4. *What is the importance of dose-concordance in establishing the biological likelihood of ozone-mediated effects occurring at relevant exposure concentrations in humans?*

   Particularly in the context of known dose information about ozone: total inhaled dose includes concentration, exposure time, and exercise duration; Hatch et al., (2013) have shown that humans and rats that are exposed to ozone at rest achieve similar alveolar ozone doses, and that humans exercising at 5-times a resting ventilation rate achieved an ~ 5-times higher alveolar ozone dose; and that ozone concentrations are 2-10 times lower indoors where people spend most of their time.

   I think this is a very important issue and one that has not been resolved or evaluated by EPA in weighing the evidence across different studies (i.e., human chamber studies and animal studies).

5. *Is there evidence that the animal models used to assess ozone effects (largely rats, mice, and non-human primates) are more, less, or similarly sensitive to ozone-mediated adverse effects compared to humans, at approximately equal inhaled doses?*

   Again, this is a valid question, and given the evidence as presented in the ISA it is difficult to answer. I think that EPA's assertion that some of the high exposure levels used in the animal studies (based on the Hatch et al., 2013 study) are relevant to ambient exposures in humans is likely to be simplistic at best, and a more detailed analysis to support an answer to this question is warranted.

Causality Question

6. *In the absence of a causality diagram to direct the choice of variables to control in an epidemiological study, how can we judge whether a study has appropriately controlled for confounders, and has not inappropriately controlled for colliders (which can open up pathways between variables that otherwise would not be connected) or mediators (and thereby controlled away the effect)?*

   I don't think this is a new issue and as noted above, this is a particularly important limitation of observational air pollution studies. The study summaries that EPA presents in the ISA fall short of identifying the various limitations in the epidemiological literature and this remains an area of weakness in the overall evaluation of ozone health effects. For determining plausible (but not necessarily absolute) causation, a full integration of all lines of evidence is necessary. As noted previously, relying on only epidemiological evidence is not sufficient.

Questions from Dr. Steven Packham

Question 1. When a causal relationship is conclusive to a high degree of scientific certainty as it is in this case, should this take precedence over causal inference when drafting a NAAQS ISA?

The human chamber studies indicate that for adults exercising for several hours and exposed to a certain concentration of ozone, small (but measurable) FEV1 decrements are observed, generally at levels above about 72 ppb. These effects are reversible and are likely not clinically significant (i.e., people may not even notice the changes in lung function). Some people do report subjective symptoms. As noted in our paper [see Goodman, JE; Prueitt, RL; Chandalia, J; Sax, SN. 2014. "Evaluation of adverse human lung function effects in controlled ozone exposure studies." *J. Appl. Toxicol.* 34(5):516-24], the results from chamber studies need to consider several areas of potential uncertainty. This includes the inter-individual variation in FEV1 measurements that can occur due to factors other than ozone exposures. For example, this intra-individual variation in FEV1 measurements can be up to about 5% and could explain some of the observed statistically significant lung function changes at lower ozone exposure levels. In addition, the level of exertion in these studies is an important consideration as the exercise regimen in most of these (i.e., 40 L min–1 for 6–8 h) is equivalent to work performed during a day of heavy manual labor common of outdoor workers. Therefore, the level of exertion and differential impact of exercise on study participants, in terms of its impact on FEV1 decrements, needs to be considered.

Question 2: Given evidence available from controlled human exposures substantiating causal relationships with a number of physiological responses, including beneficially confounding interactions of ozone on PM clearance, should Sub-section ES4.1 Health Effects in the Draft's Executive Summary, and the entire Integrated Synthesis section of the Draft be rewritten?

The points that Dr. Packham bring up are all valid and need to be incorporated in the discussions in the various relevant sections. While edits may be warranted, I don't think these sections necessarily need to be re-written.

Question 3: Looking ahead, do you think toxicology, clinical human studies, and biomedical research disciplines should be given more explicit and balanced consideration in the development of the present, and future, O3 ISAs with the objective to validate causal relationships and determine hourly inhalation dosage rates for adverse inflammatory responses in pulmonary tissues?

Toxicological, clinical (chamber) human exposure studies and biomedical (mechanistic) studies are included in the evaluation of health impacts in the ISA. In fact, as noted by Dr. Packham in the prior NAAQS assessment of ozone effects, human chamber studies played a key role in identifying the recommended level of exposure to ozone and lead to the revision of the NAAQS – in particular the study by Schelegle et al. (2009). However, there needs to be an important distinction between effects that may be small, transient and reversible (and likely not clinically significant) and effects that may be more serious. It is not possible to determine the level of exposure that would likely result in overt and irreversible harm using human chamber studies, as it would not be ethical. Therefore, the chamber studies must be interpreted within the context of other lines of evidence, including relevant exposures in animal studies and data regarding mode of action in mechanistic studies. For ozone exposures, it is particularly important (as Dr. Packham noted in the background for this question) to understand underlying defense mechanisms in the lungs that would be protective unless these mechanisms are

overwhelmed (i.e., at high doses) and can no longer provide the necessary protection against oxidative harm. I also agree with his evaluation of cumulative dose and that EPA's figures summarizing these data miss the mark with regards to cumulative dose over time. As noted above, these chamber studies involved healthy individuals exercising over several hours and this needs to be accounted for in the interpretation of the findings.

I like the idea of providing information to the public regarding levels that may results in some small reversible effects (that vary across individuals) and setting alerts that people can access on an app to obtain that information. Again, there does need to be a distinction between these benchmarks or guidelines and levels that would likely produce long-term adverse effects and this should also be communicated to the public.

Questions from Dr. James Boylan

Appendix 1 – Atmospheric Source, Chemistry, Meteorology, Trends, and Background

I have not reviewed this section is a lot of detail, but it appears that most of the section is dedicated to defining ozone background and reviewing recent studies and trends in background ozone concentrations. This is an important area of research and warrants detailed discussion. Consideration of background levels of ozone is important in making policy decisions regarding the NAAQS, particularly as the NAAQS is lowered to levels that approach background in many areas (e.g., the western US and areas influenced by wildfires).

I did find some EPA statements to be somewhat contradictory in its summary of the evidence, noting for example, on the one hand that background ozone concentrations can range from 20-50 ppb, and modeling error can be as large as 10 ppb, but then dismissing background ozone by noting that anthropogenic sources contribute to large proportions of measure ozone concentrations (at least on high ozone days). In addition, EPA concludes that the general trend of increasing background ozone from international sources has seen a slow down or reversal more recently. These statements imply that background ozone may not be or should not be given much importance. I would argue that the upper end of the background estimates is quite high relative to actual ambient levels in may areas.

Appendix 2 – Exposure to Ambient Ozone

I briefly reviewed the exposure section. While the section provided a good overview of exposure concepts and models, including new methods that are being use in epidemiological studies to refine exposure measurements, there were some things missing. In particular, I found that the section emphasized too much aspects of exposure measurement error in epidemiological studies and how they can be interpreted, without consideration of the poor correlations between indoor and ambient concentrations and personal and ambient ozone exposures given the time spent indoors. If available, EPA should not only present indoor-outdoor and personal-outdoor ratios, but an estimate of what exposures may actually be vis a vis ambient exposures. It seems to me that the issue of exposure measurement error is very different for ozone than other air pollutants (e.g., particulate matter), because of the impact of ozone scavenging on surfaces. That is, if the exposure is so small because of little

penetration of ozone indoors and no indoor sources, can we conclude that ambient measurements – even if refined by modeling – are good surrogates of exposure for people spending most of their time indoors? This is an area that warrants more discussion and consideration than what is provided currently in the ISA. It may be that for large portions of the population, possibly excluding outdoor workers or people that exercise outdoors, ozone exposures may be very low or negligible, but this information is not included in this section and should be.

Appendix 9 – The Role of Tropospheric Ozone in Climate Effects

I am not sufficiently familiar with all of the climate change processes that could be influenced by ozone to appropriately comment on these questions. However, I think this is a new and important area of research and I am glad that EPA is including an assessment of these impacts in the ISA.

Questions from Dr. Corey Masuca

Appendix 1 Atmospheric Source, Chemistry, Meteorology, Trends, and Background Ozone
1.3.1 Precursor Sources

I am not sufficiently familiar with ozone chemistry an all precursors to provide a thorough answer.

1.3.1.2.1 Global Methane

Same as above

1.3.1.2.2 International Emissions of Ozone Precursors

This section focuses on international transport of ozone precursors.

What about local/state/regional transport of ozone precursors?

I agree that this is an equally important issue that warrants consideration, particularly if areas that are exceeding the NAAQS due to transport of precursors from other states.

1.3.1.3.2 Biogenic Volatile Organic Compounds (VOCs)

I am not familiar enough with the models to provide a comment to this question.

1.4 Ozone Photochemistry

With the advent of monitoring for speciated compounds including PAMS and Near-Road Monitoring (NOy), should there be further discussions about the individual chemicals gleaned from the specialized monitoring.

If the speciated chemicals are relevant to ozone formation, they should be included in discussions regarding the impact of these chemicals on ozone.

1.5 Inter-Annual Variability and Longer Term Trends in Meteorological Effects on Anthropogenic and US Background (USB) Ozone

I agree that topography (e.g., in Los Angeles) is an important consideration in the accumulation of air pollution. This should be included with examples. I am unsure about the independent effects due to relative humidity.

Appendix 2 Exposure to Ambient Ozone

2.3 Exposure Assessment Methods

While monitoring, including fixed, ambient monitors and personal and microenvironmental monitors are highlighted, what about remote sensing? Biological sampling in blood or tissue?

This is a good question, although I do not have much to contribute in terms of specific answers.

2.3.2.1 Spatial Interpolation

While attempting to quantity concentrations at locations and areas between concentration points is included under 2.3.2 Modeling, many of these exact same methods (i.e., data averaging, IDW, and kriging) are also utilized for Monitoring data shortcomings.
I agree – these are common methods used when data are limited in general.

2.4.1 Time-Activity Data

Is it possible that ozone exposure through time-activity data may be reduced due to temperature alone, as more people tend to avoid time spent outdoors in the summers during extremely warm/hot/humid, stagnant days which are oftentimes conditions for greater ozone formation?

Yes, I believe this is very possible. I have only seen data on people that may alter their behavior if they have asthma (or kids with asthma) on high pollution days. I have not seen data on activity patterns altered by temperature alone, although these data may be available.

Miscellaneous Question(s)

Due to exposure to ozone being disproportionate for disparate (i.e., lower income, children), should this be emphasis in this section, in lieu of regression analysis confounding/covariate in epidemiological studies for low(er) SES?

Yes – I agree that some discussion on the potential disparate exposures in low income communities or children would be warranted. Although unlike particulate matter, because ozone is formed over time from precursors, it is not always the case that lower income communities might have the highest exposures. EPA does discuss time-activity patterns in children that might contribute to higher exposures

(i.e., because of greater time spent outdoors), but this could be made more explicit (e.g., are there any exposure data or actual exposure measurements to confirm this?).

Dr. Duncan Thomas, University of Southern California

Questions from Dr. Tony Cox

1. *Can valid determinations of manipulative or interventional causation … be made based on observed associations of the types analyzed in the ISA?*

Response: I discussed this question extensively in my earlier comments on the draft $PM_{2.5}$ PA. Briefly, there exist methods of causal inference designed to reanalyze observational epidemiology data as if it were from a randomized trial with the goal of estimating the Average Causal Effect of a hypothesized intervention. There have been a handful of such publications in the air pollution literature (see my previous citations), as well as a few "Accountability" studies based on real world "natural experiments" (work stoppages, introduction of new regulations, short term interventions around Olympic Games, etc.). However, the vast bulk of air pollution studies have not been designed or analyzed for the purpose of assessing manipulative or interventional causation. Nevertheless, the consistency of the findings from numerous observational studies, the concordance with short term human experimental studies (e.g., chamber or panel studies), and animal experiments, along with other lines of evidence supporting biological plausibility, as outlined in the preface to the ISA, allows a causal interpretation in terms of the likely effect of air pollution on the various health endpoints, if not a quantitative estimate of the predicted magnitude of the effect of a hypothetical intervention. See also my responses to some parts of question 2.

This point has been cogently discussed in a recent commentary by Carone, Dominici & Sheppard (2019), who conclude "In our view, causal inference methods should not be used as another opportunity to weaponize science against itself. Policymakers cannot wait for the data, study designs, and analytic tools that will ensure unarguable causal inferences: stalling until perfect evidence arises is irresponsible and does not protect public health." See also (Goldman and Dominici 2019): "a requirement of manipulative causation fails to recognize the full depth and robustness of existing approaches in epidemiology, statistics, and causal inference and the degree to which they deal with confounding factors."

2. *The following questions are intended to help assess the conceptual clarity and meaning of the causal determination categories, and of key conclusions expressed using them, such as those in Table ES-1 (p. E-5) of the Draft ISA*
    a. *Is this actually a "formal causal framework"?*

Response: That depends upon the meaning of the term in quotes. The approach used in the ISA does not exploit the emerging framework of "causal inference" that constitutes one type of "formal causal framework." However, the "weight of evidence" machinery (Committee_to_Review_the_IRIS_Process 2014) used here is certainly a well-established and appropriate formal framework for reaching causal judgements combining evidence across scientific disciplines. The machinery of statistical causal inference is not capable of or intended to synthesize evidence across multiple studies from multiple scientific disciplines.

Unlike Dr. Cox, I do not find the definitions in Table II of the Preface (and their application in Tables IS-4 and IS-5) to be "logically incoherent and ambiguous"; rather, the definitions are operational rather than conceptual, in line with those used by IARC and other scientific agencies, based on specific criteria for the types of evidence required to attain each category, e.g., for a determination of a "causal relationship" the following is required:

> "… Generally, the determination is based on multiple studies conducted by multiple research groups, and evidence that is considered sufficient to infer a causal relationship is usually obtained from the joint consideration of many lines of evidence that reinforce each other. "

b. Does the ISA's causal determination framework clearly distinguish between necessary and sufficient causation?

Response: see my response to question *h* below.

c. does a "causal relationship" determination imply a manipulative causal relationship?

Response: That depends upon the context in which the term "causal relationship" is used. In the statistical literature on causal inference, yes, the goal is to estimate the effect of an intervention on differences in the expected outcome within an individual under different hypothetical scenarios. In the epidemiological literature — and as used in the ISA — it refers to the existence of a mechanism under which exposure is a contributing factor, which may imply that a change in exposure would be expected to change the outcome, but that is not the primary sense in which the term is used. See also my response to question 2.*h*.

d. Can causal determinations be incorrect?

Response: Yes, of course, any human judgment could be incorrect. That is true of those reached by a large body of experts as well but is much less likely!

e. is it clear how uncertainty about which category is correct should be (or has been) resolved in assigning a final causal determination category?

Response: While the process for deciding upon which category is appropriate is clearly described, I do not see much if any discussion about any disagreements about the choice of category were resolved, except in terms of justification for changes in the categorization since the 2013 ISA (Tables IS-4 and IS-5). These explanations seem cogent to me.

f. is it clear how observations could be used to test and falsify a given causal determination if it is not correct?

Response: The question appears to ask whether "relevant data" not already considered, i.e., new studies, could falsify a conclusion in the ISA. While it is always possible that new data will

emerge that leads one to question a previous determination, such speculation would be beyond the scope of the ISA.

g. is the correctness of each causal determination in table ES-1 formally and transparently evaluated in the ISA?

Response: I find that the evidence provided in section IS.4 and the supporting appendices to provide compelling support for the determinations in Table ES-1 and the supporting Appendices (to the extent that I have been able to read parts of it and to the extent of my epidemiologic expertise), and the process for reaching these judgments to be clearly described and transparent.

h. Does a determination that an exposure-response (or concentration-response (C-R)) relationship is a "causal relationship" imply that it is entirely causal,

Response: No, but that depends upon what is meant by "entirely causal." Epidemiologists have long recognized a "complex web of causation (MacMahon and Pugh 1970) meaning that no single factor is ever both necessary and sufficient to cause disease. A "causal relationship" is generally held to mean that a risk factor is a real component of one of the "sufficient component causes" of disease (Rothman 1976).

i. Does a determination that a C-R relationship is a "causal relationship" imply 100% certainty that it is causal?

Response: Obviously that depends upon the confidence with which that judgment has been reached. This seems like a semantic quibble.

j. Does a determination that a C-R relationship is a "causal relationship" imply that it is causal for every member of a population,

Response: not necessarily. More likely the magnitude of the effect will vary across subgroups of the population, but biology being essentially the same across all humans, it is likely that a causal association in the population at large will be true to some extent for any subgroup. Of course, it is possible that some subgroups will have no association at all: e.g., men are not likely to be at risk of ovarian cancer and those who lack a particular genotype that is essential for metabolism of a particular agent may be absolutely immune.

k. Are the five categories mutually exclusive?

Response: yes.

l. Are the five categories collectively exhaustive?

Response: yes.

m. Can a body of evidence be categorized as "likely to be causal" if the probability of causality based on the evidence is less than 50%?

Response: Causal inference methods aim to estimate the "Average Causal Effect", not the probability of causality. The "Probability of Causation" (PC) is an estimate of the probability that a specific individual's disease was caused by some aspect of his exposure history, essentially an individualized version of the epidemiologic concept Population Attributable Risk Fraction. The PC has been frequently used in toxic tort litigation and setting guidelines for compensation policy, although it has come in for criticism, but is irrelevant for judging the causality of an observational association in populations.

3. *… are its conclusions derived by valid inference from true premises? Are the stated conclusions implied by the data and analyses used to support them? Are they consistent with other data and analyses that are at least as good as those selected? Are they appropriately caveated?*

    a. *Study selection and interpretation:*

Response: The various questions for subsection 3(*a):i-vi* and (*b-g*) below require extensive substance matter knowledge of air pollution epidemiology and toxicology that are beyond my expertise. I have to defer to the EPA experts who drafted this ISA and the other consultants to respond to the specifics about the selection and interpretation of specific studies and any omissions therein. However, I have added a few comments on specific questions about which I have some expertise.

i. *Is it clear that the ISA's study selection process has successfully provided a comprehensive, trustworthy, and unbiased selection of the best available science on ozone and health effects?*
ii. *Is it clear why results from Moore (2008) are included and cited as "key evidence" but contrary results from Moore (2012) are excluded? More generally, is it clear that study inclusion and exclusion criteria were applied systematically and neutrally to identify and select the best and most up-to-date studies to inform the ISA's conclusions?*

Response: I agree that the Moore (2012) should be discussed. These two papers were among the few air pollution publications that used causal inference techniques, as I cited in my $PM_{2.5}$ draft PA comments, so the apparent difference between the conclusions of the two papers merits comment. It appears that the lack of significance from the later paper may be due in part from having to restrict to the 41% of the original 195 geographical grids for which the "experimental treatment assignment (ETA)" assumption was valid (i.e., that the probability of exposure assignment being above or below the 90 ppb threshold being analyzed was between 10% and 90%). The authors go on (following the sentence quoted by Dr. Cox) to state: "the fact that the CMRIER analysis does not provide significant results may be due to the lack of power to detect an effect with inverse weighting estimation. A more efficient estimation approach like TMLE estimation [10] could improve the estimation precision."

iii. *Are there other studies that are omitted from the ISA that should be included?*
iv. *Are there studies included in the ISA that should be omitted?*

v. *Is it clear that the process followed in selecting and summarizing scientific studies in the ISA was sufficient to assure accurate, unbiased, up-to-date, and trustworthy summaries of the relevant scientific literature to inform causal determination judgments?*

vi. *Do you find in the Executive Summary a clear explanation of the extent to which the key evidence supporting the ISA's causal determinations consists of, is sensitive to, or is derived from unverified modeling assumptions, or from modeling assumptions that more recent literature has found to be incorrect or inadequate? Have you found information in the ISA on sensitivity of causal determination conclusions to untested, uncertain, or incorrect assumptions?*

b. *Were the epidemiological studies used to support the causal determinations summarized in Table ES-1 (p. ES-5) and Figure ES-2 (p. ES-6) appropriately designed and analyzed to provide valid scientific information and valid causal conclusions about effects of possible future interventions (rather than just conclusions about historical statistical associations)?*

c. *Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" adequately controlled for potential confounding and residual confounding by variables such as income and weather variables**?***

Response: Of those studies cited in support of these determinations that I am familiar with, the authors have gone to appropriate lengths to control for such confounders, to the extent possible with the available data. Of course, residual confounding can never be excluded from any observational epidemiology study.

d. *Is it clear that the individual studies cited in support of the ISA's causal determinations of "causal" or "likely to be causal" have adequately controlled for biases due to exposure estimation errors or exposure misclassification errors?*

Response: While the various studies differ in their methods of exposure estimation, few to my knowledge have used formal methods of measurement error correction. However, the bias from measurement error would generally be in the direction of reducing effect sizes and power, not introducing false positives. An exception would be for multi-pollutant analyses where it is possible for some effect of a better measured and causal pollutant to be transferred to the estimate for a worse measured noncausal one, as discussed at length in my response to the same question in my comments on the draft $PM_{2.5}$ PA.

e. *Do you find in the Executive Summary, or elsewhere in the ISA, a clear explanation of the extent to which the key evidence supporting the ISA's causal determinations is sensitive to uncontrolled or incompletely controlled confounding and/or ecological associations?*

f. *More generally, is it clear how criteria for individual study quality were applied to each study used in making causal determinations, and what the results were? (See Table Annex 6-1, cf p. 6-67.) Is it clear how the limitations of each individual study were taken into account in causally interpreting their reported associations and in making causal determinations?*

g. *Does the ISA make clear how its causal determinations would change if evidence from associations caused by confounding, residual confounding, measurement error, or unverified modeling assumptions were excluded?*

Response: This seems rather speculative, absent any evidence that those studies included have failed to adequately address these possible biases.

4. *Is the biological evidence presented in the ISA to support causal determinations correctly stated, correctly interpreted, relevant for predicting effects of changes in the ozone NAAQS, and up-to-date?*

Response: Yes, as explained in my responses to various other questions by Dr. Cox. (By "biological evidence" I assume you mean to include epidemiology, amongst the other lines of evidence, e.g., toxicology, which would be largely beyond my expertise.) See the following response in particular for the question about "*predicting effects of changes in the ozone NAAQS*".

5. *Does the biological evidence presented in the ISA provide well-validated scientific information suitable for predicting the effects on public health of changing NAAQS standard for ozone?*

Response: I believe the evidence presented in the ISA is suitable for reaching a causal interpretation of the effects of air pollution on human health. The ISA does not address the implications of potential changes in the NAAQS; it is my understanding that that will be addressed in the draft Policy Assessment document that I have not seen yet.

6. *Is each of the causal determinations summarized in Table ES-1 (especially those labeled "causal relationship" or "likely to be causal relationship") the only possible causal determination conclusion that is justified by, or consistent, with current scientific evidence? Could different causal determinations be equally well justified (or better justified) by the information presented, or by the totality of current scientific evidence?*

Response: Any judgment of a "causal" or "likely to be causal" relationship is potentially subject to differences of opinion amongst experts. It is my opinion that the various determinations summarized in Table ES-1 are well justified by the totality of the evidence, based on what I have read in sections IS.4, the supporting appendices, and my general background knowledge of the field of air pollution epidemiology. I do not claim to have read more than a portion of the ISA or to have an exhaustive knowledge of the substance matter of air pollution epidemiology, however. That said, I would consider it highly unlikely that any other conclusions would be "equally well justified" or "better justified" than those reached by the authors of the draft ISA.

7. *Are there changes in the design, analysis, selection, or interpretation of individual studies or in the ISA's processes for interpreting and summarizing them that would improve the validity, credibility, and transparency of the ISA's scientific reasoning and conclusions?*

Response: Obviously I would welcome wider application of the techniques of causal inference to observational studies, along the lines of those publications I cited in my response to the draft PM2.5 PA. That said, I believe that the weight of evidence approach used by EPA to evaluate the totality of the

evidence, experimental and observational, to be highly appropriate and I have no further suggestions for improvement in that process.

Questions from Dr. Mark Frampton

1. ***Change in causality determination for short-term cardiovascular effects since the 2013 ISA.*** *Please comment on the strengths and weaknesses of the epidemiology literature with regard to CV effects of short-term ozone exposure. Are there key studies that are missing? Are the remaining weaknesses, along with the other new evidence, sufficient to justify the change in causality determination?*

Response: While I find the rationale provided in the ISA for this change in causality determination to be compelling, I do not have the comprehensive knowledge of the air pollution epidemiology literature to address whether there are key studies missing, other new evidence, or inadequately considered weaknesses.

2. ***Metabolic effects, new determination of "likely" for both short- and long-term exposure.*** *Is there sufficient epidemiological evidence of metabolic effects to justify the "likely" determination for both short- and long-term exposures? Are there additional studies that should be considered?*

Response: Same answer as for the previous question. From my limited experience collaborating with my epidemiologic colleagues on air pollution and obesity, metabolic syndrome, and related conditions, it certainly seems this is a "hot topic" and that "likely" is not unreasonable.

3. ***Change in causality determination for total mortality since the 2013 ISA.*** *Please comment on the strengths and weaknesses of the epidemiology literature with regard to short-term ozone exposure and total mortality. Are there key studies that are missing? Does the available evidence justify the change in causality determination for total mortality? Also please note that, for effects with causal or likely causal determination, the EPA has restricted consideration of epidemiological studies to those in North America (see PECOS Tool, section 6.1.1.1, page 6-3). That was the case for this determination. Are there epidemiological studies of mortality outside of North America that should be considered?*

Response: I was surprised by the reclassification of the association of short-term ozone exposure and mortality from "likely" to "suggestive," given the enormous body of time-series studies, including large multi-city studies (e.g., NMMAPS and APHEA) showing relationships with cause-specific and total mortality. There is no discussion of the rationale for this change in section ES-4 of the executive decision. Section IS-4.3.1 discusses this evidence only briefly in terms of respiratory mortality only and does not really provide a rationale for downgrading this association. Indeed, Table IS-4 highlights the abundant evidence support the "likely" classification in the 2013 ISA and then says for the 2019 one

> "Recent epidemiologic evidence for respiratory mortality is limited, but there remains evidence of consistent, positive associations, specifically in the summer months, with mean daily 8-h max ozone concentrations between **8.7 and 63 ppb.** When recent

evidence is considered in the context of the larger number of studies evaluated in the 2013 Ozone ISA, there remains consistent evidence of an association between short-term ozone exposure and respiratory mortality." [emphasis added]

I have no idea why the EPA would have a policy of relying only on studies from North America. This does not make sense to me, given the availability of high-quality studies from Europe (e.g., the aforementioned APHEA) and elsewhere.

Questions from Dr. Corey Masuca

None of Dr. Masuca's question pertaining to Appendices 1 and 2 are within my area of expertise. I will confine my response to the following one.

Miscellaneous Question(s): Due to exposure to ozone being disproportionate for disparate (i.e., lower income, children), should this be emphasis in a this section, in lieu of regression analysis confounding/covariate in epidemiological studies for low(er) SES?

Response: It is certainly true that exposure is disproportionately distributed, a serious concern known as "environmental justice." This also renders socioeconomic and other factors associated with exposure to be confounders requiring control for epidemiological associations. That ozone affects underprivileged communities disproportionately is worth pointing out in this section but does not alter the need to discuss the appropriateness of the methods used for confounder control. It is my opinion, based on those studies I have been involved in myself or have read in the literature, that the vast majority of the studies relied upon in the ISA have addressed this issue appropriately, to the extent possible with the available data.

Questions from Dr. Sabine Lange

1. *It has been established that associations found in an epidemiology study can be due to: causation, bias, chance, and/or confounding. If the concept of statistical significance is not useful in epidemiology studies, then how do the study authors/EPA rule out that chance has caused the observed association?*

Response: In addition to bias (of which confounding is one kind), chance can certainly lead to non-causal associations. Assessment of statistical significance is essential to judge the likelihood that an association could be due to chance, so it's incorrect to say that it "is not useful in epidemiology studies." Despite the longstanding and on-going debates about the usefulness specifically of *p*-values for this purpose (Greenland et al. 2016, Wasserstein and Lazar 2016), as opposed to a variety of other approaches (e.g., confidence intervals, Bayes Factors, etc.), they remain the most commonly used method for judging the possibility of chance. I do not see that the EPA has dismissed statistical significance testing in its evaluation of the evidence, although they correctly do incorporate "trends in data and reproducibility of results" as well as other considerations in their evaluation of the epidemiologic evidence.

2. *Am I correct in understanding that the intention of ozone case-crossover studies is to compare the ozone concentrations on a day when a health effect occurred for a person, to the ozone concentrations on a day when that health effect did not occur for that person?*

Response: Yes, that is the correct interpretation. An advantage of this design is that by making comparisons with an individual, between-individual confounding is completely eliminated, as are any factors that do not vary over time. While factors other than pollution that do vary over time, like weather, could still be confounders, these can be controlled in the analysis by standard statistical adjustment methods, as in case-control or time-series studies.

3. *If so, then it would be important that some other factor (not related to ozone) did not prevent the health event from occurring on a control day. These studies often use days before and after the health event as control days, but for mortality studies (such as Di et al., 2017), how can a day after death be used as a control day? It doesn't matter what the ozone concentrations are after a person's death, that person would not be able to respond to that concentration. How should we interpret case-crossover studies that use control days after the event (particularly mortality) occurred?*

Response: The original case-control design (Maclure 1991) involved a comparison of exposure at the time of the event (or some pre-specified time prior to it to allow for lag effects) to that at some previous comparable ("referent") time. For example, the referent time might involve the same day of the week to control for systematic weekly variation in pollution levels and/or confounders. My colleague, Bill Navidi (1998) pointed out, however, that seasonal variation and especially long-term trends in pollution levels could lead to bias if referent times always preceded event times, even if one or more entire year cycles were included; while there would be no bias if there were no long-term trends and if pollution followed a perfectly symmetric (e.g., sinusoidal) seasonal pattern, departures from such symmetry, as are common for both pollution and meteorology, would lead to bias. Instead, he proposed the "bidirectional case-crossover" design, in which two referent times, one before and one after, equally spaced around the event time, are used. The original Macluer design was intended to study personal time-varying characteristics such as behaviors that could be "triggers" for an event like death or heart attack; in this setting, it would be impossible to observe a behavior that occurred after death! In air pollution studies, however, personal behaviors are not being studied, but ambient exposures are and these can be measured and used meaningfully for comparison. While it is obviously true that pollution *after* the event could not be causally related to the event, the purpose of this design is to get an unbiased estimate of the expected exposure at the time of the event for comparison with the actual exposure at that time and can be interpreted as a sampling-based analog of the standard time-series approach for acute effects (Bateson and Schwartz 1999, Fung et al. 2003, Lu and Zeger 2007). Various versions of this design have subsequently been widely adopted in air pollution studies. Although the original bidirectional design has subsequently been shown to be slightly biased (Lumley and Levy 2000), a modified version involving using fixed time-strata, comparing exposures at event times within each stratum with those at all or selected times (e.g., day-of-week matched times) within the same stratum before and after the event, has been shown to be unbiased (Levy et al. 2001a, Janes et al. 2005a, Janes et al. 2005b), and this design has become the standard in substantive studies (e.g., (Levy et al. 2001b, Di et al. 2017)). As Mittleman (2005) says, "this strategy should be considered the de facto standard approach to the analysis of data arising in studies of the short-term effect of air pollution and weather" (see also references therein for additional studies using this design).

4. *What is the importance of dose-concordance in establishing the biological likelihood of ozone-mediated effects occurring at relevant exposure concentrations in humans?*

Response: If by "dose-concordance" you mean comparability of doses to animals and humans from similar external concentrations and ventilation rates, I would expect that there are so many factors that differ that it would be unreasonable to expect the same dose-response relationships, even if doses could be scaled in comparable units.

5. *Is there evidence that the animal models used to assess ozone effects (largely rats, mice, and non-human primates) are more, less, or similarly sensitive to ozone-mediated adverse effects compared to humans, at approximately equal inhaled doses?*

Response: See my response to the previous question. Not being a toxicologist, however, these questions are largely beyond my expertise.

6. *In the absence of a causality diagram to direct the choice of variables to control in an epidemiological study, how can we judge whether a study has appropriately controlled for confounders, and has not inappropriately controlled for colliders (which can open up pathways between variables that otherwise would not be connected) or mediators (and thereby controlled away the effect)?*

Response: Very good question! Directed Acyclic Graphs (DAGs) can be useful tools for visualizing hypothetical relationships among observed and latent variables and for structuring an appropriate analysis strategy (Greenland et al. 1999). Investigators typically have such pictures in mind when conducting an analysis, although they are seldom presented formally in a substance matter publication (they are more commonly included in statistical methods papers). The basic principles that confounders must be controlled using the best available data on known risk factors (or surrogates for unmeasured factors in an attempt to minimize residual confounding), and that intermediate variables on a causal pathway from exposure to disease not be adjusted for, nor for colliders (that are determined by both exposure and disease but are not causal for disease), are well understood. The art is in deciding which variables are or are not appropriate to adjust for. While there are a variety of formal statistical methods for dealing with adjustment uncertainty (Maldonado and Greenland 1993, Greenland 1996, Viallefont et al. 2001, Crainiceanu et al. 2007, Pope and Burnett 2007), it remains a matter for expert judgment, both by the original investigators and by critical readers.

Questions for Non-CASAC member Consultants on the Ozone ISA from Dr. Steven Packham

1. *When a causal relationship is conclusive to a high degree of scientific certainty as it is in this case, should this take precedence over causal inference when drafting a NAAQS ISA?*

Response: Yes. In my opinion, while formal statistical causal inference methods have a useful role, particularly for evaluating the predicted effect of a hypothesized intervention, evaluation of whether a health effect of air pollution is causal requires a synthesis of evidence from multiple types of studies, which goes far beyond what these methods are capable of. I support the general weight of evidence framework used by EPA in this and other ISAs for this purpose.

2. *Given evidence available from controlled human exposures substantiating causal relationships with a number of physiological responses, including beneficially confounding interactions of ozone on PM clearance, should Sub-section ES4.1 Health Effects in the Draft's Executive Summary, and the entire Integrated Synthesis section of the Draft be rewritten?*

Response: I see no mention of PM interactions with ozone in section ES4.1. To the extent that there is compelling evidence the ozone enhances PM clearance and mitigates its adverse effects (literature I am unfamiliar with), then it would seem appropriate to mention that here.

3. *Looking ahead, do you think toxicology, clinical human studies, and biomedical research disciplines should be given more explicit and balanced consideration in the development of the present, and future, O3 ISAs with the objective to validate causal relationships and determine hourly inhalation dosage rates for adverse inflammatory responses in pulmonary tissues?*

Response: Not being an expert in toxicology, clinical medicine, or biomedical research, all I can say is that I believe *all* these disciplines are relevant, as is my field of epidemiology, which in my opinion has the most direct relevance to human morbidity and mortality. Whether they deserve "*more* explicit and balanced consideration" implies that they are given inadequate consideration in the present and maybe future ISA, which I am not really qualified to answer. It does seem to me that the present draft ISA has attempted to assess all the relevant information from the various disciplines and incorporate them appropriately in their weight of evidence framework, at least with respect to the causality of the various relationships. I can't comment specifically on the comparability of dose rates between humans and model systems or their implications for pulmonary responses, other than to reiterate that there are many factors that differ among them other than dose that could make such comparisons dubious. Dr. Packham's comment that "ozone-induced FEV1 effects are temporary, reversible, and occur at a lower inhaled dose than a truly adverse health effect" sounds plausible, but beyond my expertise to critique.

Questions from Dr. James Boylan

His questions regarding Appendices 1 and 5 and the first few of those on Appendix 2 are beyond my expertise. I will, however, respond to the following two regarding Appendix 2:

- *Is the discussion on copollutant correlations and potential for confounding (Section 2.5) accurate and complete? If not, what additional information needs to be included?*
- *Is the discussion on interpreting exposure measurement error for use in epidemiology studies (Section 2.6) accurate and complete? If not, what additional information needs to be included?*

Response: I found both of these sections to be very clear, accurate, comprehensive, and well reasoned. I am not aware of any additional information that needs to be included. In particular, I agree with the conclusions in section 2.7 about the likely direction and magnitude of any biases introduced by co-pollutant correlations and measurement error. While section 2.6 could elaborate slightly on the available techniques for correcting for exposure measurement error (Thomas et al. 1993, Carroll et al. 2006), these have been applied only rarely in substantive epidemiology studies, so that would really be necessary.

REFERENCES

Bateson, T. F. and J. Schwartz (1999). "Control for seasonal variation and time trend in case-crossover studies of acute effects of environmental exposures." Epidemiology **10**(5): 539-544.

Carone, M., F. Dominici and L. Sheppard (2019). "In Pursuit of Evidence in Air Pollution Epidemiology: The Role of Causally Driven Data Science." Epidemiology **Publish Ahead of Print**.

Carroll, R. J., D. Ruppert, L. A. Stefanski and C. M. Crainiceanu (2006). Measurement Error in Nonlinear Models: A Modern Perspective (2nd Ed.). London, Chapman and Hall CRC Press.

Committee_to_Review_the_IRIS_Process (2014). Review of EPA's Integrated Risk Information System (IRIS) Process. Washington DC, National Academy of Sciences.

Crainiceanu, C. M., F. Dominici and G. Parmigiani (2007). "Adjustment uncertainty in effect estimation." Biometrika **95**(3): 635-651.

Di, Q., L. Dai, Y. Wang, A. Zanobetti, C. Choirat, J. D. Schwartz and F. Dominici (2017). "Association of Short-term Exposure to Air Pollution With Mortality in Older Adults." JAMA **318**(24): 2446-2456.

Fung, K. Y., D. Krewski, Y. Chen, R. Burnett and S. Cakmak (2003). "Comparison of time series and case-crossover analyses of air pollution and hospital admission data." Int J Epidemiol **32**(6): 1064-1070.

Goldman, G. T. and F. Dominici (2019). "Don't abandon evidence and process on air pollution policy." Science **363**(6434): 1398-1400.

Greenland, S. (1996). "Basic methods for sensitivity analysis of biases." Int J Epidemiol **25**(6): 1107-1116.

Greenland, S., J. Pearl and J. M. Robins (1999). "Causal diagrams for epidemiologic research." Epidemiology **10**(1): 37-48.

Greenland, S., S. J. Senn, K. J. Rothman, J. B. Carlin, C. Poole, S. N. Goodman and D. G. Altman (2016). "Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations." Eur J Epidemiol **31**(4): 337-350.

Janes, H., L. Sheppard and T. Lumley (2005a). "Case-crossover analyses of air pollution exposure data: referent selection strategies and their implications for bias." Epidemiology **16**(6): 717-726.

Janes, H., L. Sheppard and T. Lumley (2005b). "Overlap bias in the case-crossover design, with application to air pollution exposures." Stat Med **24**(2): 285-300.

Levy, D., T. Lumley, L. Sheppard, J. Kaufman and H. Checkoway (2001a). "Referent selection in case-crossover analyses of acute health effects of air pollution." Epidemiology **12**(2): 186-192.

Levy, D., L. Sheppard, H. Checkoway, J. Kaufman, T. Lumley, J. Koenig and D. Siscovick (2001b). "A case-crossover analysis of particulate matter air pollution and out-of-hospital primary cardiac arrest." Epidemiology **12**(2): 193-199.

Lu, Y. and S. L. Zeger (2007). "On the equivalence of case-crossover and time series methods in environmental epidemiology." Biostatistics **8**(2): 337-344.

Lumley, T. and D. Levy (2000). "Bias in the case – crossover design: implications for studies of air pollution." Environmetrics **11**(6): 689-704.

Maclure, M. (1991). "The case-crossover design: a method for studying transient effects on the risk of acute events." Am J Epidemiol **133**(2): 144-153.

MacMahon, B. and T. F. Pugh (1970). Epidemiology: principles and methods. Boston, Little, Brown and Co.

Maldonado, G. and S. Greenland (1993). "Simulation study of confounder-selection strategies." Am J Epidemiol **138**(11): 923-936.

Mittleman, M. A. (2005). "Optimal referent selection strategies in case-crossover studies: a settled issue." Epidemiology **16**(6): 715-716.

Navidi, W. (1998). "Bidirectional case-crossover designs for exposures with time trends." Biometrics **54**: 596-605.

Pope, C. A., 3rd and R. T. Burnett (2007). "Confounding in air pollution epidemiology: the broader context." Epidemiology **18**(4): 424-426; discussion 427-428.

Rothman, K. J. (1976). "Causes." American Journal of Epidemiology **104**(6): 587-592.

Thomas, D. C., D. Stram and J. Dwyer (1993). "Exposure measurement error: Influence on exposure-disease relationships and methods of correction." Ann Rev Publ Health **14**: 69-93.

Viallefont, V., A. E. Raftery and S. Richardson (2001). "Variable selection and Bayesian model averaging in case-control studies." Stat Med **20**(21): 3215-3230.

Wasserstein, R. L. and N. A. Lazar (2016). "The ASA Statement on p-Values: Context, Process, and Purpose." The American Statistician **70**(2): 129-133.